



UNIVERSIDAD
NACIONAL
AUTÓNOMA DE
NICARAGUA,
MANAGUA
UNAN-MANAGUA



Congreso
**Científico de
Educación**

Persona · Familia · Comunidad

Revista

Lengua y Literatura

EDICION ESPECIAL II

**Validación de la encuesta de pertinencia para el diseño
curricular de posgrado en Medicina Veterinaria**

EJE TEMÁTICO:

La gestión y calidad en educación

ISSN: 2707-0107
Vol. 12/ Edición Especial
2026



Validación de la encuesta de pertinencia para el diseño curricular de posgrado en Medicina Veterinaria

Validation of the relevance survey for postgraduate curriculum design in veterinary medicine

José Miguel Lara Lazo
jlara0591@gmail.com

<https://orcid.org/0009-0007-3604-6966>

Universidad Nacional Autónoma de Nicaragua, Managua

César Arévalo Cuadra
carevalo@unan.edu.ni

<https://orcid.org/0000-0001-8954-3586>

Universidad Nacional Autónoma de Nicaragua, Managua

EJE TEMÁTICO: La gestión y calidad en educación

© UNAN-Managua

Recibido: noviembre 2025 Aprobado: febrero 2026

<https://doi.org/10.5377/rl.v12iEspecial2.22476>



RESUMEN

Este trabajo tuvo como propósito garantizar la validez de contenido y la confiabilidad del juicio experto del proceso de validación del instrumento INSP001, diseñado para determinar la pertinencia del proyecto de diseño curricular para la carrera de posgrado Maestría en Medicina Interna Veterinaria, bajo el modelo por competencias de la UNAN-Managua; con un enfoque cuantitativo-descriptivo, La validación fue realizada por nueve expertos, seleccionados bajo criterios de experiencia académica y grado de formación (maestría y doctorado), los que analizaron tres criterios de evaluación (claridad, coherencia y pertinencia) a nueve ítems mediante una escala tipo Likert de cinco puntos. Los datos se analizaron mediante indicadores psicométricos como el índice de Validez de contenido por ítem (I-CVI), índice de validez de contenido promedio (S-CVI/Ave), Kappa de Polit modificada y coeficiente V de Aiken, mediante el uso de la herramienta de inteligencia artificial Julius AI. Los resultados evidenciaron un alto nivel de acuerdo entre los expertos, con valores de I-CVI entre 0.90 y 1.00, S-CVI/Ave superiores a 0.94, Kappa entre 0.89 y 1.00 y Aiken V en el rango de 0.77 a 0.97, lo cual confirma la validez de contenido excelente y confiabilidad del instrumento. Los ítems con valores ligeramente inferiores fueron conservados por su relevancia conceptual, realizándose ajustes semánticos menores.

Palabras Claves:

Confiabilidad, Evaluación Curricular, Pertinencia, Posgrado, Validez.

ABSTRACT

The purpose of this study was to ensure the content validity and reliability of the expert judgment in the validation process of the INSP001 instrument. This instrument was designed to determine the relevance of the curriculum design project for the postgraduate program Master's in Veterinary Internal Medicine, based on the competency-based model of UNAN-Managua, utilizing a quantitative-descriptive approach. The validation was carried out by nine experts, selected based on criteria of academic experience and level of education (Master's and Doctoral degrees). They analyzed three evaluation criteria (clarity, coherence, and relevance) for nine items using a five-point Likert scale. The data were analyzed using psychometric indicators, including the Item Content Validity Index (I-CVI), the Scale Content Validity Index/Average (S-CVI/Ave), Modified Polit's Kappa, and Aiken's V coefficient, with the assistance of the artificial intelligence tool Julius AI. The results showed a high level of agreement among the experts, with I-CVI values ranging from 0.90 to 1.00, S-CVI/Ave values exceeding 0.94, Kappa between 0.89 and 1.00, and Aiken's V in the range of 0.77 to 0.97. This confirms the excellent content validity and reliability of the instrument. Items with slightly lower values were retained due to their conceptual relevance, and minor semantic adjustments were made.

Keywords: reliability, curricular evaluation, relevance, postgraduate, validity.

INTRODUCCIÓN

La universidad genera conocimiento mediante la investigación (Salas y Mayorga, 2023); y la investigación es el eje principal para el avance y desarrollo de cualquier disciplina. (Vizcaíno et al, 2023). La pertinencia de los programas universitarios se relaciona directamente con la dinámica social (Salas y Mayorga, 2023); por lo que al momento de generar propuestas innovadoras estas deben de ser vinculantes con el entorno socioeconómico donde se espera desarrollar.

Según Corral de Franco (2022), la relevancia y confiabilidad de los resultados, es la brújula de una investigación científica y académica; por lo que se vuelve imperativo el contar con instrumentos bien contruidos que garanticen la pertinencia, calidad y veracidad de la información recolectada. Lo que permite indicar que los instrumentos son una parte fundamental de todas las investigaciones, razón por la cual López et al (2019) afirmaron que «la validación de instrumentos es considerada, por el alcance de su rigor científico, un tipo de estudio con sus propias características y procedimientos». (p. 441)

Villasís et al (2018), expuso que el término «validez» se refería «a lo que es verdadero o lo que se acerca a la verdad». Asimismo, Reveles (2020), afirmó que «la validez, es un término usado para denotar cualidades siendo una de ellas el grado en que una explicación es aceptada para ser generalizable» (p. 1). Por lo que se puede indicar que los resultados de un estudio se consideran confiables cuando el instrumento mediante el cual fueron recolectados posee un alto grado de validez.

Es por esto que, el diseño, validación e implementación de herramientas metodológicas de recolección de datos es una etapa fundamental para asegurar la calidad, validez y confiabilidad de los resultados (Reveles, 2020).

Este proceso les permite a las universidades desarrollar su capital intelectual de manera pertinente lo que indica un alto grado de vinculación con la sociedad.

Para Villasís et al (2018), todos los instrumentos de recolección de datos como lo son: cuestionarios, encuestas o rúbricas, deben de cumplir con criterios estrictos que garanticen su relevación en función de los objetivos establecidos y su capacidad para medir con precisión las variables en estudio. Ya que sin este proceso se corre el riesgo de obtener datos inconsistentes o ambiguos, afectando así la confiabilidad de los hallazgos (Salas y Mayorga, 2023)

MATERIAL Y MÉTODOS

La presente investigación es de carácter cuantitativo, descriptivo, no experimental y de corte transversal. Está orientada al análisis de calidad del instrumento de encuesta de pertinencia (INSP001) diseñado como parte del proyecto del diseño curricular para la carrera de posgrado Maestría en Medicina Interna Veterinaria, bajo el modelo por competencias de la UNAN-Managua.

Esta investigación adoptó el método hermenéutico y analítico, dado que se centró en la interpretación y valoración crítica de los ítems por parte de un validador (especialistas en el área), considerando los criterios de claridad, coherencia y pertinencia.

Tabla 1.

Dimensiones evaluadas por cada ítem del instrumento

Criterio	Definición operacional	Indicadores de valoración
Claridad	Grado en que el ítem está redactado de manera comprensible, sin ambigüedades ni confusión.	Uso adecuado del lenguaje, precisión semántica, facilidad de comprensión por parte del lector.
Coherencia	Relación lógica y consistencia entre el ítem, la dimensión que mide y los objetivos del estudio.	Correspondencia con la variable de interés y alineación con el marco teórico.
Pertinencia	Relevancia y necesidad del ítem dentro del instrumento para medir el constructo en estudio.	Grado en que el ítem aporta información útil, esencial y no redundante.

El instrumento Encuesta de Pertinencia (INSP001) fue evaluado por un panel de nueve expertos que aceptaron participar de forma voluntaria (muestra no probabilística e intencional), con formación académica a nivel de maestría y doctorado, y con experiencia en educación superior y curricular. Con el objetivo de mantener la confidencialidad de los evaluadores, cada uno de ellos fue codificado utilizando el acrónimo ESP (especialista), acompañado de la letra D si el evaluador contaba con doctorado (ESPD) o con la letra M si contaba con maestría (ESPM), seguido por un número consecutivo asignado según orden de registro (ESPD001/ESPM001).

Tabla 2.

Caracterización de los expertos evaluadores

N°	Grado	Área	IES	País	Experiencia	Codificación
1	PhD	Educación Superior	UNAN León	NI	13 años	ESPD001
2	PhD	Doctor en Ciencias	Universidad de Zaragoza	ES	26 años	ESPD002
3	MSc	Gerencia de servicios públicos agropecuarios	UCA	NI	21 años	ESPM002
4	MSc	Aprendizaje, Cognición y Desarrollo Educativo.	UNIR	MX	16 años	ESPM003
5	MSc	Docencia Universitaria	Universidad Paulo Freire	NI	19 años	ESPM004
6	MSc	Higiene e inocuidad de los alimentos	UNAN León	NI	5 años	ESPM005
7	MSc	Formulación y Evaluación de Proyectos	Universidad Rubén Darío	NI	10 años	ESPM006

8	MSc	Sanidad Animal de Pequeñas Especies, con Mención en Medicina Interna	UNAN León	NI	11 años	ESPM007
9	MSc	Currículo Universitario	UPOLI	NI	15 años	ESPM008

Nota: NI=Nicaragua, ES=España, GT=Guatemala, MX=México.

Asimismo, se empleó un sistema de codificación único para el instrumento, permitiendo identificarlo de manera segura, y facilitando la trazabilidad en el análisis de los datos. La codificación se construyó combinando tres elementos principales:

- **Tipo de instrumento:** se utilizan abreviaturas que identifican el tipo de análisis.
- **Numeración correlativa:** cada instrumento recibe un número progresivo (001, 002, 003), lo que garantiza un registro único
- **Actor o enfoque:** en el caso de las encuestas, se especifica a qué población está dirigida.

Tabla 3.

Codificación de instrumentos

Instrumento	Codificación	Objetivo	Dirigido a
Pertinencia	INSP001	Diseñado para determinar la pertinencia de la propuesta de la carrera de posgrado.	Profesionales de la medicina veterinaria

Tal y como se puede apreciar en la tabla 4, el instrumento utilizado permitió a los expertos evaluar a cada ítem o pregunta perteneciente al instrumento las tres dimensiones, utilizando una escala tipo Likert de uno a cinco.

Tabla 4.

Ejemplificación del instrumento de validación

Ítem del instrumento	Claridad ¿Está redactado de forma comprensible?	Coherencia ¿Se ajusta al objetivo de la encuesta?	Pertinencia ¿Aporta información útil para evaluar la pertinencia?	Observaciones del expertp
Pregunta 1				

Nota: cada uno de los instrumentos utilizados por los validadores analizó cada una de las preguntas o ítems, de cada documento metodológico.

El proceso de validación dio inicio con el envío de la carta solicitando la participación del experto como evaluador, posterior a la aceptación, se envió el instrumento de validación en un archivo de Word digital a cada uno de forma individual manteniendo la confidencialidad y anonimato de los participantes.

Una vez que los validadores entregaron los datos, estos fueron consolidados en una base de datos tipo tabla utilizando el programa Microsoft Excel, para facilitar su procesamiento y el análisis mediante estadística descriptiva (promedios). Asimismo, se realizó análisis estadístico mediante el uso de Julius AI (<https://julius.ai/>), el cual es una herramienta de inteligencia artificial, diseñada para el análisis de datos, mediante el uso de modelos de lenguaje de texto extenso (LLM) y generación de códigos de análisis en Python. (Murillo y López, 2025); la cual aplicó verificación, limpieza, análisis descriptivo y cálculos psicométrico.

Tabla 5.

Procedimientos realizados mediante Julius AI

Etapa	Descripción	Herramientas	Dimensión evaluada	Objetivo
1. Preparación de datos	Revisión de completitud, detección de valores perdidos y coherencia de respuestas.	Base de datos en Excel Julius IA		Ordenamiento de la base de datos
3. Cálculo del I-CVI	Proporción de expertos que calificaron un ítem ≥ 4 ("acuerdo relevante").	Fórmula: $I-CVI = \frac{n_{agree}}{N_{total}}$	Validez de contenido por ítem	Validez
4. Cálculo del S-CVI/Ave	Promedio de I-CVI por criterio, como medida global de validez de contenido.	Julius IA	Validez global por criterio	Validez
5. Cálculo del Kappa modificado (Polit)	Ajuste del I-CVI para compensar el acuerdo por azar.	Fórmula: $K = \frac{(I-CVI - P_c)}{(1 - P_c)}$	Consistencias del juicio	Confiabilidad
6. Cálculo del V de Aiken	Coefficiente que mide la relevancia y adecuación de cada ítem considerando la escala ordinal.	Fórmula: $V = \frac{\sum s}{[n(c-1)]}$	Intensidad del acuerdo ordinal	Confiabilidad (complementaria)

EJE TEMÁTICO: La gestión y calidad en educación

Según Roebianto et al (2023), el CVI, permite cuantificar el grado de acuerdo entre expertos sobre la relevancia de los ítems. El I-CVI, representa la proporción de jueces (validadores) que consideran un ítem esencial (valor ≥ 4 en la escala), mientras que el S-CVI/Ave constituye el promedio de los I-CVI de todos los ítems de una dimensión o del instrumento completo.

Al momento de la interpretación se tomó que un I-CVI ≥ 0.78 como aceptable, mientras que para S-CVI/Ave ≥ 0.90 se tomó como validez global aceptable. (Castillo et al, 2024)

Asimismo, se calculó el índice de Kappa modificado de Polit, el cual corrige el valor del I-CVI considerando el acuerdo que podría ocurrir por azar (Orts, 2011).

Al momento de la interpretación de los resultados se tomó a $K > 0.74$ como excelente, mientras que los rangos de 0.60-0.74 se tomó como bueno, y < 0.60 como pobre. (Polit et al, 2007); los cuales se calcularon mediante la fórmula:

$$K = \frac{I - CVI - P_c}{1 - P_c}$$

Donde $P_c = \frac{N!}{A!(N-A)!} \times 0.5^n$, Siendo A el número de expertos en acuerdo y N el total de expertos

También, se determinó el coeficiente V de Aiken, el cual mide la intensidad de la validez de contenido. (Merino, 2023), considerando la escala ordinal empleada por los expertos. El valor de V varía entre 0 y 1; cuanto más cercano a 1, mayor es la validez, por lo que valores ≥ 0.80 se interpretó como muy adecuados. (Penfield y Giacobbi, 2004); realizando el cálculo mediante la fórmula:

$$V = \frac{\sum s}{n(c-1)}$$

Donde $s = x - l_0$, siendo x la puntuación otorgada al ítem, l_0 el valor mínimo de la escala, n el número de jueces y c el número de categorías.

RESULTADOS Y DISCUSIÓN

El instrumento INSP001, corresponde al instrumento para determinar la pertinencia del proyecto de diseño curricular para la carrera de posgrado Maestría en Medicina Interna Veterinaria, bajo el modelo por competencias de la UNAN-Managua. Este instrumento será dirigido a profesionales de la Medicina Veterinaria y consta de nueve ítems (preguntas).

Tabla 6.

Validación del INSP001 por ítem

Instrumento		INSP001									ESP/Criterio	Global por especialista	
Ítems		1	2	3	4	5	6	7	8	9			
Codificación de los evaluadores	ESPD001	Claridad	4	4	4	4	4	4	4	4	4	4.00	4.00
		Coherencia	4	4	4	4	4	4	4	4	4	4.00	
		Pertinencia	4	4	4	4	4	4	4	4	4	4.00	
	ESPD002	Claridad	5	5	5	5	5	5	5	5	5	5.00	5.00
		Coherencia	5	5	5	5	5	5	5	5	5	5.00	
		Pertinencia	5	5	5	5	5	5	5	5	5	5.00	
	ESPM002	Claridad	5	5	5	5	5	5	5	5	3	4.78	4.78
		Coherencia	5	5	5	5	5	5	5	5	3	4.78	
		Pertinencia	5	5	5	5	5	5	5	5	3	4.78	
	ESPM003	Claridad	5	5	5	5	5	5	5	5	5	5.00	5.00
		Coherencia	5	5	5	5	5	5	5	5	5	5.00	
		Pertinencia	5	5	5	5	5	5	5	5	5	5.00	
	ESPM004	Claridad	4	5	4	5	5	4	4	5	5	4.56	4.56
		Coherencia	4	5	4	5	5	4	4	5	5	4.56	
		Pertinencia	4	5	4	5	5	4	4	5	5	4.56	
	ESPM005	Claridad	5	5	4	5	5	5	5	5	5	4.89	4.96
		Coherencia	5	5	5	5	5	5	5	5	5	5.00	
		Pertinencia	5	5	5	5	5	5	5	5	5	5.00	
	ESPM006	Claridad	5	5	4	5	5	3	3	5	5	4.44	4.56
		Coherencia	5	5	4	5	5	3	3	5	5	4.44	
		Pertinencia	5	5	4	5	5	4	5	5	5	4.78	
	ESPM007	Claridad	5	5	4	5	5	5	5	5	5	4.89	4.96
		Coherencia	5	5	5	5	5	5	5	5	5	5.00	
		Pertinencia	5	5	5	5	5	5	5	5	5	5.00	
ESPM008	Claridad	4	5	2	5	3	4	5	5	5	4.22	4.22	
	Coherencia	4	5	2	5	3	4	5	5	5	4.22		
	Pertinencia	4	5	2	5	3	4	5	5	5	4.22		

Nota: se observa las calificaciones otorgadas por los evaluadores a cada uno de los ítems del INSP001, el color rojo indica ítems con calificación igual o menor a 3.

EJE TEMÁTICO: La gestión y calidad en educación

El análisis evidenció que el instrumento INSP001 (Pertinencia) obtuvo, en general, una valoración positiva por parte de los nueve expertos participantes. La mayoría de los ítems fueron calificados en los rangos de 4 (de acuerdo) y 5 (totalmente de acuerdo) en los criterios de claridad, coherencia y pertinencia. Esto indica que, en términos generales, los enunciados presentan un nivel adecuado de comprensión y se ajustan al objetivo de medir la pertinencia de la propuesta de maestría.

No obstante, se identificaron algunos ítems con menor consenso, reflejado en puntuaciones de 3 (ni de acuerdo ni en desacuerdo) y, en algunos casos, de 2 (en desacuerdo). En particular, el ítem 3 recibió tres valoraciones de 2, lo que sugiere que algunos expertos percibieron deficiencias en su formulación, ya sea en claridad o en su relación con el constructo evaluado. Asimismo, los ítems 6, 7 y 9 obtuvieron valoraciones de 3, evidenciando dudas respecto a su coherencia; lo que resalta la necesidad de realizar ajustes específicos en estos ítems, con el fin de fortalecer la validez.

Por otro lado, los ítems 1, 2, 4, 5 y 8 muestran una casi unanimidad en las valoraciones altas (4 y 5), lo que confirma su solidez conceptual y la pertinencia de mantenerlos sin modificaciones. Estos resultados permiten indicar que, el instrumento de pertinencia es válido y confiable en términos generales, pero requiere ajustes en los ítems específicos 3, 5, 6, 7 y 9, para alcanzar un nivel óptimo de validez de contenido y garantizar su utilidad en el proceso de evaluación académica.

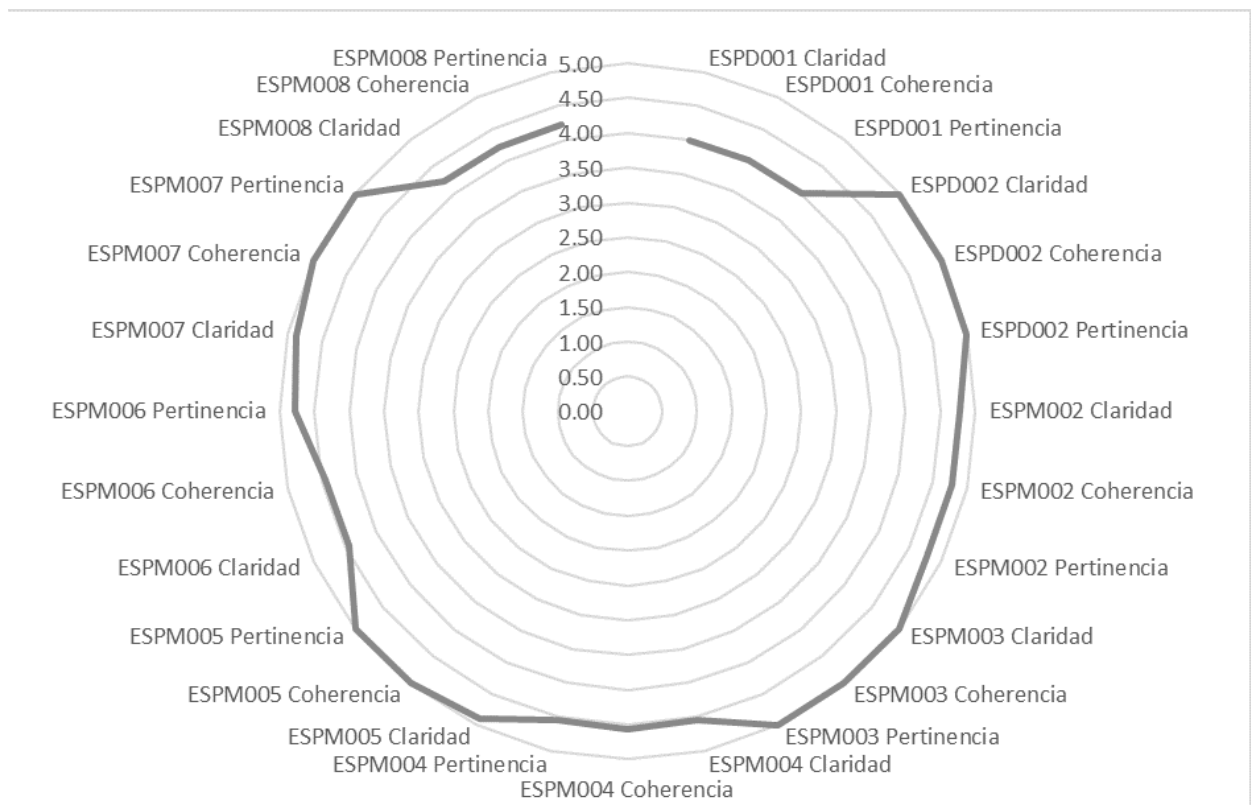
Sin embargo, a pesar de esto, se puede apreciar que el promedio de cada criterio correspondiente al INSP001, resultante de las evaluaciones emitidas por cada uno de los evaluadores, se mantienen entre 4 y 5.

El análisis del INSP001 evidenció que los valores obtenidos en los tres criterios evaluados (claridad, coherencia y pertinencia) son consistentemente altos, con promedios que oscilan entre 4 y 5. Estos resultados indican una evaluación favorable y consistente por parte de los expertos.

Para el criterio de claridad los evaluadores ESPD002 y ESPM003 calificaron al INSP001 con un 5 en la escala de Likert, seguido por ESPM005 y ESPM007 con una calificación de 4.89, por su parte los especialistas ESPM002, ESPM004, ESPM006 y ESPM008 calificaron con 4.78, 4.56, 4.44 y 4.22 respectivamente. Siendo únicamente el ESPD001 que brindó la calificación más baja con un 4.

Figura 1.

Validación promedio por Criterio del INSP001



EJE TEMÁTICO: La gestión y calidad en educación

Tal y como se aprecia en la figura 1, para el criterio de coherencia los evaluadores ESPD002, ESPM003, ESPM005 y ESPM007 concordaron con una calificación de 5, seguidamente los evaluadores ESPM002, ESPM004 y ESPM006 brindaron calificaciones de 4.78, 4.56 y 4.44 respectivamente, por su parte el ESPM008 calificó la coherencia con 4.22, mientras que el ESPD001 calificó con un 4.

Estos hallazgos son coherentes con los criterios metodológicos de validación establecidos por Polit et al (2007), quienes señalaron que la consistencia entre los ítems y el constructo es un indicador esencial para la validez.

En lo que respecta al criterio de pertinencia, los evaluadores ESPD002, ESPM003, ESPM005 y ESPM007 indicaron una calificación de 5, seguido de ESPM002 y ESPM006 quienes concordaron con una calificación de 4.78; por su parte los ESPM004 y ESPM008 calificaron con 4.56 y 4.22 respectivamente, siendo el ESPD001 la única calificación de 4.

Los resultados anteriores son congruentes con la validación global del INSP001.

Tabla 7.

Valoración global del INSP001

Criterios	Claridad	Coherencia	Pertinencia
Global por Criterios	4.641975309	4.666666667	4.703703704
Global del instrumento		4.670781893	

Fuente: cálculo realizado usando Julius AI (2025)

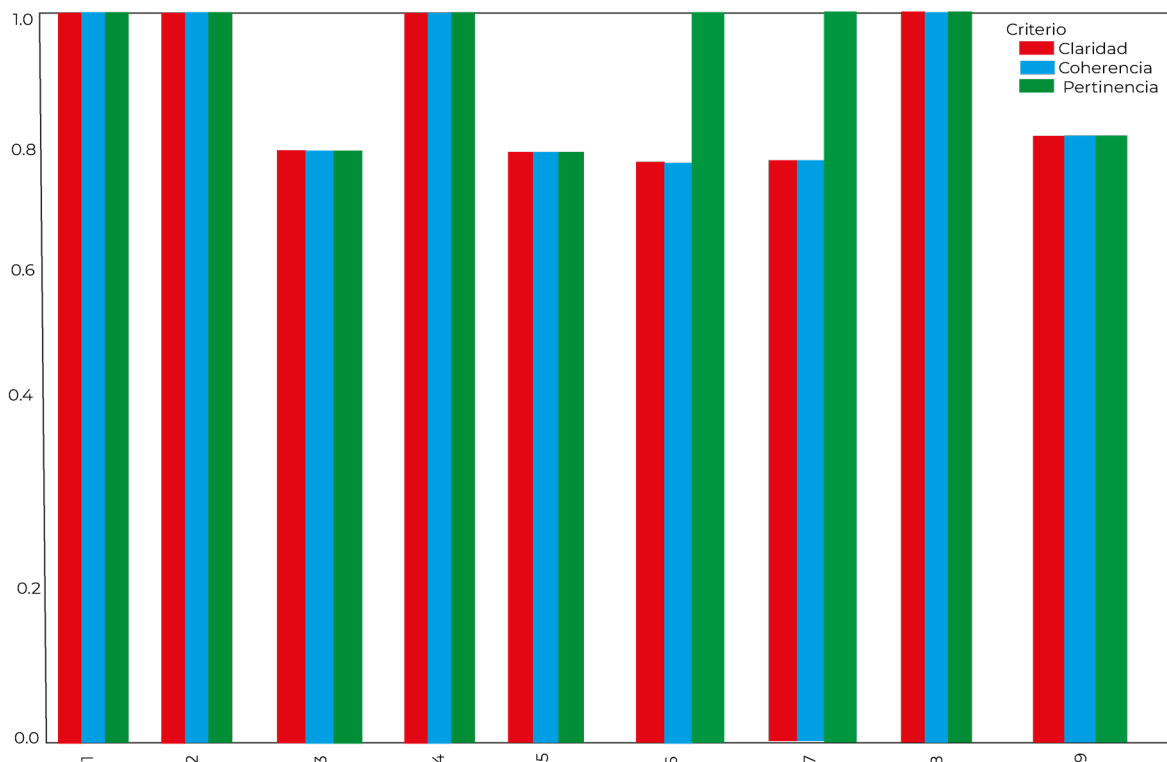
Los promedios globales del instrumento en los tres criterios fueron: claridad 4.64, coherencia 4.67 y pertinencia 4.70, lo cual se traduce en una media general superior al 4.6 y una desviación estándar baja ($\approx 0.55-0.61$).

Esto refleja un nivel alto de consenso por parte de los expertos. Lo que indica que, en general, el instrumento de pertinencia fue percibido como comprensibles, congruentes con su objetivo y adecuado para medir los aspectos relacionados con la pertinencia de la propuesta de posgrado.

En cuanto a los resultados de la validez del juicio emitido por los expertos, los resultados fueron plasmados en la figura 2, a continuación.

Figura 2.

Valores del I-CVI por ítem



Nota: ilustración de los valores del índice de validez de contenido por ítem para los criterios evaluados por los expertos. Fuente: Julius AI (2025)

EJE TEMÁTICO: La gestión y calidad en educación

Los ítems 1, 2, 3, 4, 7 y 8 presentaron un I-CVI igual a 1, en los tres criterios, lo cual indica que todos los expertos coincidieron en que dichos ítems son claros, coherentes y pertinentes. En el caso de los ítems 3, 5, 6 y 9, el valor del I-CVI desciende a 0.90, lo que quiere decir que el 90% de los expertos otorgó una calificación ≥ 4 . Sin embargo, estos valores se mantienen por encima del parámetro establecido de 0.78 indicado por Castillo et al (2024).

Para el criterio de claridad tanto como la coherencia, se presentaron valores que oscilaron entre 0.9 y 1, indicando que la redacción y comprensión, al igual que la lógica y consistencia conceptual de los ítems fue adecuada, y que solamente aquellos ítems con valores de 0.9 podría requerir un ajuste semántico menor.

En lo que corresponde al criterio de pertinencia, se pudo apreciar que fue el que presentó mayor número de ítems con I-CVI de 1, lo que refleja que, para los expertos los ítems son relevantes y alineados con la finalidad del instrumento.

Estos resultados demuestran que los expertos consideran los ítems claros, coherentes y pertinentes, con valores de I-CVI superiores al estándar mínimo; apreciándose un acuerdo inter experto elevado, lo que respalda la validez de contenido del instrumento y la fiabilidad del juicio experto, avalando así, el uso del mismo en las etapas siguientes de la investigación.

Por su parte el S-CVI/Ave, muestra la medida global de la validez de contenido, calculada a partir del promedio del I-CVI obtenido por ítem. Los resultados se aprecian en la tabla 8.

Tabla 8.
Índice de Validez de contenido global del INSP001

Criterio	S-CVI/Ave	Interpretación
Claridad	0.94	Muy alta validez de contenido
Coherencia	0.94	
Pertinencia	0.97	

Fuente: cálculo realizado usando Julius AI (2025)

Para el criterio de claridad, se indica un valor del 0.94, lo que permite indicar que el 94% de los expertos coincidieron al momento de considerar los ítems claros y comprensibles, lo que permite a su vez afirmar que el lenguaje, la estructura y la redacción de dichos ítems permiten una interpretación unívoca.

Asimismo, el criterio de coherencia refleja una alta correspondencia lógica entre cada ítem y la dimensión teórica que pretende evaluar. Con una valoración de 0.94, se puede afirmar que el 94% de los expertos perciben que los ítems se articulan correctamente con los objetivos del instrumento.

En lo que respecta a la pertinencia, se aprecia que fue el valor más alto obtenido con un 0.97. lo cual significa que el 97% de los expertos consideran que los ítems son relevantes para medir la pertinencia de la propuesta de posgrado.

Tomando como base los resultados anteriores, se puede afirmar que el instrumento cuenta con una validez sustantiva, ya que como se apreció en la tabla 8, los valores de los tres criterios superan el parámetro de ≥ 0.90 indicado por Castillo et al (2024).

La similitud en los valores obtenidos en los criterios de claridad y coherencia indica que las preguntas están bien formuladas y alineadas conceptualmente; asimismo, el aún más alto valor obtenido en el criterio de pertinencia demuestra que los expertos reconocen la relevancia de los ítems.

En cuanto a la confiabilidad de los resultados del proceso de validación del instrumento INSP001, los resultados fueron expuestos en la tabla 9.

Tabla 9.

Valores del Criterio Kappa Polit modificada y Aiken V

Ítem	Criterio	Kappa Polit modificada	Aiken V
1	Claridad	1	0.91
	Coherencia	1	0.91
	Pertinencia	1	0.91
2	Claridad	1	0.97
	Coherencia	1	0.97
	Pertinencia	1	0.97

3	Claridad	0.89	0.77
	Coherencia	0.89	0.83
	Pertinencia	0.89	0.83
4	Claridad	1	0.97
	Coherencia	1	0.97
	Pertinencia	1	0.97
5	Claridad	0.89	0.91
	Coherencia	0.89	0.91
	Pertinencia	0.89	0.91
6	Claridad	0.89	0.86
	Coherencia	0.89	0.86
	Pertinencia	1	0.88
7	Claridad	0.89	0.88
	Coherencia	0.89	0.88
	Pertinencia	1	0.94

EJE TEMÁTICO: La gestión y calidad en educación

8	Claridad	1	0.97
	Coherencia	1	0.97
	Pertinencia	1	0.97
9	Claridad	0.89	0.91
	Coherencia	0.89	0.9
	Pertinencia	0.89	0.91

Fuente: Fuente: cálculo realizado usando Julius AI (2025)

Para el criterio de claridad los valores de Kappa se encuentran entre el 0.89 y el 1, mientras que los valores de V de Aiken oscilan entre 0.77 y 0.97, indicando que existe consenso entre los s expertos, respecto a la redacción y comprensión de los ítems.

En cuanto al ítem 3, que obtuvo la valoración más baja en la V de Aiken, se podría atribuir esta valoración a diferencias de interpretación o al nivel de especificidad de la redacción, sin afectar la validez general del instrumento.

En cuanto a coherencia, los valores de Kappa igual que a claridad los valores se encontraron entre el 0.89 y el 1; mientras que, la V de Aiken mostro promedios de 0.90. Estos resultados demuestran una consistencia lógica fuerte, lo que significa que los ítems están correctamente alineados con las dimensiones teóricas del instrumento.

Por su parte, el criterio de pertinencia obtuvo los valores más altos de Kappa con ítems mayormente en 1 y V de Aiken con la mayoría de los ítems en 0.97; lo que indica el máximo acuerdo entre los expertos referente a la relevancia de los mismos.

Los valores en el análisis de kappa de polit modificado, presentan rangos que van del 0.89 al 1, valores que afirman que los expertos encontraron los ítems relevantes (calificaciones ≥ 4). Los valores obtenidos superan el parámetro de ≥ 0.74 establecido por Polit et al (2007). En cuanto al V de Aiken, todos los ítems presentaron valores superiores al parámetro de 0.75 de Penfield y Giacobbi (2004), lo que indica una intensidad de acuerdo sólida.

CONCLUSIONES

Los resultados obtenidos evidencian un sólido acuerdo interjueces entre los tres criterios evaluados (claridad, coherencia y pertinencia), con valores de I-CVI entre 0.90 y 1, S-CVI/Ave superiores a 0.94, Kappa de Polit modificada entre 0.98 y 1 y coeficiente V de Aiken en el rango de 0.77 a 0.97. Estos resultados permiten confirmar que el instrumento INSP001, presenta validez de contenido excelente y un acuerdo interevaluadores confiable, afirmando así que los ítems son representativos, comprensibles y pertinentes para el contexto del diseño de la propuesta del posgrado

Asimismo, se afirma que, los ítems con valores ligeramente menores no comprometen la estructura conceptual del instrumento, sino que sugieren ajustes semánticos mínimos orientados a mejorar la redacción y precisión de los enunciados.

Es por lo anterior que se concluye que el INSP001 cumple con los estándares de validación y confiabilidad, lo que lo convierte en una herramienta metodológica sólida para el desarrollo de la investigación futura que se pretende.

REFERENCIAS

- Carral de Franco, Y. (2022). Validez y confiabilidad en instrumentos de investigación: una mirada teórica . Revista Ciencias de la Educación, XXXII(60), 562-586. <https://www.servicio.bc.uc.edu.ve/educacion/revista/60/art06.pdf>
- Castillo, P., Gaecía, R., Hernández, J., Poblano, E., y Noriega, S. (2024). Validez de contenido de un instrumento de medición para medir el éxito en la implementación del TPM. Revista Iberoamericana para la investigación y el desarrollo educativo, XIV(28). doi:10.23913/ride.v14i28.1928
- López, R., Avello, R., Palmero, D., Sánchez, S., y Quintana, M. (2019). Validación de instrumentos como garantía de la credibilidad en las investigaciones científicas . Revista Cubana de Medicina Militar, II(48), 441-450. <http://scielo.sld.cu/pdf/mil/v48s1/1561-3046-mil-48-s1-e390.pdf>
- Merino, C. (2023). Coeficientes V de Aiken: diferencias en los juicios de validez de contenido. MHSalud, XX(1), 1-10. doi:10.15359/mhs.20-1.3
- Murillo, D., y López, S. (2025). Julius.ia : Herramienta IA generativa para el análisis y visualización de datos. Panamá: Universidad Tecnológica de Panamá. https://www.researchgate.net/publication/389232584_Juliusia_Herramienta_IA_generativa_para_el analisis_y_visualizacion_de_datos
- Orts, M. (2011). Validez de contenido del practice environment scale of the nursing work index (PES-NWI) en el ámbito europeo. España: Universidad de Alicante. <https://rua.ua.es/server/api/core/bitstreams/f7edf1d6-14ff-437f-bfe2-ddbae5218a9d/content>
- Penfield, R., y Giacobbi, P. (2004). Applying a Score Confidence Interval to Aiken's Item Content-Relevance Index. Measurement in Physical Education and Exercise Science, IV(8), 213-225. doi: 10.1207/s15327841mpee0804_3

- Polit, D., Beck, C., y Owen, S. (2007). Is the CVI an Acceptable Indicator of Content Validity? Appraisal and Recommendations. *Research in Nursing y Health*, XXX, 459-467. doi:10.1002/nur.20199
- Reveles, L. (2020). Validez de los instrumentos de medición. (UNAM) Obtenido de Unidades de Apoyo para el Aprendizaje. https://repositorio-uapa.cuaed.unam.mx/repositorio/moodle/pluginfile.php/2142/mod_resource/content/1/Contenido/index.html
- Roebianto, A., Savitri, I., Aulia, I., Suciyan, A., y Mubarokah, L. (Marzo de 2023). Content validity: Definition and procedure of content validation in psychological research. *TPM*, XXX(1), 5-18. doi:10.4473/TPM30.1.1
- Salas, M., y Mayorga, A. (2023). La calidad y pertinencia del posgrado en diversos contextos. Una revisión literaria. *Revista latinoamericana de ciencias sociales y humanidades*, IV(2), 1219. <https://latam.redilat.org/index.php/lt/article/view/679/882>
- Villasís, M., Márquez, H., Zurita, J., Miranda, G., y Escamilla, A. (2018). El protocolo de investigación VII. Validez y confiabilidad de las mediciones. *Revista Alergia México*, IV(65), 414-421. <https://www.scielo.org.mx/pdf/ram/v65n4/2448-9190-ram-65-04-414.pdf>
- Vizcaíno, P., Maldonado, I., y Cedeño, R. (2023). Metodología de la investigación científica: guía práctica. *Ciencia Latina Revista Científica Multidisciplinar*, VII(4), 9723-9762. doi:https://doi.org/10.37811/cl_rcm.v7i4.7658