



Fecha de presentación: septiembre, 2025 Fecha de aceptación: noviembre, 2025 Fecha de publicación: enero, 2026

4

Validación de instrumentos en modelo de desempeño profesoral en ciencias médicas

Validation of instruments in a model of professorial performance in medical sciences

M. Sc. Yataria Abreu Vázquez ¹
sanyataria@infomed.sld.cu

ORCID: <https://orcid.org/0000-0001-6864-5591>



Lic. Ricardo Fernández Rodríguez ²
ricarfr@infomed.sld.cu

ORCID: <https://orcid.org/0000-0003-3318-4021>



Lic. René Arenas Gutiérrez ³
renearenasg@gmail.com

ORCID: <https://orcid.org/0000-0002-6784-713X>



Cita sugerida (APA, séptima edición)

Abreu, Y., Fernández, R. y René Arenas, R. (2026). Validación de instrumentos en modelo de desempeño profesoral en ciencias médicas. *Revista Mapa*, 4(42), 53-65.

<http://revistamapa.org/index.php/es>

1Facultad de Ciencias Médicas "Julio Trigo López", Universidad de Ciencias Médicas de La Habana, Cuba.

2 Facultad de Ciencias Médicas "Julio Trigo López", Universidad de Ciencias Médicas de La Habana, Cuba.

3 Facultad de Ciencias Médicas "Julio Trigo López", Universidad de Ciencias Médicas de La Habana, Cuba.

RESUMEN

Objetivo: Validar por expertos la dimensión "Uso de las TIC para la investigación" como componente de un modelo para el mejoramiento del desempeño científico-metodológico e investigativo del profesorado de una facultad de ciencias médicas. **Método:** Se realizó un estudio metodológico de validación de contenido. La variable se operacionalizó en cinco dimensiones de la dirección educacional (planificación, ejecución, evaluación, coordinación y asesoramiento), generando 15 ítems iniciales. Un panel de 10 expertos evaluó la pertinencia de los ítems. Tras una reunión de conciliación para resolver discrepancias, se elaboró una versión definitiva de 13 ítems para la prueba de desempeño, la cual fue evaluada por otro panel de expertos aplicando los criterios de Moriyama. Se calcularon el Índice de Validez de Contenido (S-CVI), el V de Aiken, el Kappa de Fleiss y el Alfa de Cronbach. **Resultados:** La versión final mostró una validez de contenido excelente (S-CVI/Ave = 0.983; V de Aiken promedio = 0.992). La concordancia interjueces fue casi perfecta (κ de Fleiss = 0.825). El Alfa de Cronbach, interpretado con cautela, fue de 0.72. **Conclusión:** La variable "Uso de las TIC para la investigación" fue operacionalizada y validada con éxito. Los instrumentos derivados (prueba de desempeño y encuesta a directivos) poseen propiedades psicométricas sólidas que garantizan su idoneidad para la aplicación en el contexto de la educación médica superior.

Palabras Clave: desempeño docente, educación médica, investigación educativa, método Delphi, validez de contenido

ABSTRACT

Objective: To validate by experts the dimension "Use of ICT for Research" as a component of a model for improving the scientific-methodological and research performance of professors in a faculty of medical sciences. **Methods:** A methodological content validation study was conducted. The variable was operationalized into five dimensions of educational management (planning, execution, evaluation, coordination, and advising), generating 15 initial items. A panel of 10 experts evaluated the relevance of the items. After a conciliation meeting to resolve discrepancies, a definitive version of 13 items was developed for the performance test, which was evaluated by another panel of experts applying the Moriyama criteria. The Scale-level Content Validity Index (S-CVI), Aiken's V, Fleiss' Kappa, and Cronbach's Alpha were calculated. **Results:** The final version showed excellent content validity (S-CVI/Ave = 0.983; Average Aiken's V = 0.992). Inter-judge agreement was almost perfect (Fleiss's κ = 0.825). Cronbach's Alpha, interpreted with caution, was 0.72. **Conclusion:** The variable "Use of ICT for Research" was successfully operationalized and validated. The derived instruments (performance test and managerial survey) possess solid psychometric properties that guarantee their suitability for application in the context of higher medical education.

keywords: teaching performance, medical education, educational research, Delphi method, content validity

INTRODUCCIÓN

La integración efectiva de las Tecnologías de la Información y la Comunicación (TIC) en el proceso investigativo constituye un factor determinante del desempeño docente en la educación superior contemporánea (Barrios Coronel et al., 2017). Estas tecnologías no solo son herramientas de apoyo, sino elementos transformadores que “aceleran la generación del conocimiento y fomentan la colaboración global” (Barrios Coronel et al., 2017, p. 3. Para evaluar y mejorar este desempeño, es fundamental contar con instrumentos de medición válidos y confiables.

La operacionalización de constructos complejos, como el “uso de las TIC para la investigación”, es un paso metodológico esencial que permite traducir conceptos teóricos en indicadores observables y medibles (Añorga Morales, 2014). Una operacionalización sólida debe sustentarse en dimensiones teóricamente coherentes. En este estudio, se adoptaron las cinco funciones de la dirección educacional propuestas por Valle Lima et al. (2015): planificación, ejecución, evaluación, coordinación y asesoramiento, ofreciendo un marco integral para evaluar el desempeño investigativo mediado por TIC.

Sin embargo, la operacionalización teórica requiere una validación empírica que garantice la calidad de los instrumentos. La validación por expertos es una técnica consolidada para evaluar la validez de contenido, juzgando la pertinencia, claridad, coherencia y relevancia de los ítems (Escobar-Pérez & Cuervo-Martínez, 2008). La aplicación de criterios estandarizados, como los propuestos por Moriyama, permite una evaluación sistemática y rigurosa.

Este artículo presenta el proceso integral de operacionalización y validación de la variable “Uso de las TIC para la investigación” dentro de un proyecto más amplio destinado a mejorar el desempeño docente. El objetivo fue validar un

55

Yataria Abreu Vázquez

Ricardo Fernández Rodríguez

René Arenas Gutiérrez

conjunto de ítems que, integrados en una prueba de desempeño y una encuesta para directivos, midieran de forma fiable y válida este constructo multidimensional.

MÉTODOS

Se realizó un estudio metodológico para la validación de contenido de instrumentos, siguiendo un proceso secuencial de dos fases: (1) operacionalización teórica de la variable y diseño inicial de ítems, y (2) validación por dos paneles de expertos mediante juicio cuantitativo y cualitativo.

Operacionalización de la Variable

La variable “Uso de las TIC para la investigación” se definió operacionalmente como “el conjunto de técnicas y dispositivos empleados para el tratamiento y la transmisión de datos, que constituyen una herramienta fundamental en la investigación” (1). Esta variable se desglosó en cinco dimensiones, alineadas con las funciones de la dirección educacional (3):

1. Planificación: Uso de TIC para consultar fuentes bibliográficas y planificar recursos investigativos.
2. Ejecución: Uso de TIC para el procesamiento de datos, elaboración de comunicaciones, presentaciones y otros recursos durante la investigación.
3. Evaluación: Uso de TIC para el procesamiento estadístico y la valoración de los recursos producidos.
4. Coordinación: Frecuencia de la coordinación interna y externa para el uso de las TIC en investigación.
5. Asesoramiento: Frecuencia con que se ofrece y se recibe asesoramiento sobre el uso de TIC en investigación.

A partir de este marco, se diseñó una matriz de operacionalización con 15 ítems iniciales.

Procedimiento de Validación por Expertos

- Panel de Expertos: Se conformó un panel de 10 especialistas (6 Doctores en Ciencias y 4 Másters) con experiencia en metodología de la investigación, posgrado e informática, garantizando una evaluación multidisciplinar.
- Fase 1 - Evaluación de Pertinencia: Los expertos evaluaron los 15 ítems iniciales con los criterios “Mantener”, “Eliminar” o “Modificar”. Se estableció a priori un criterio de consenso predeterminada: mantener (100% de acuerdo), reanalizar (70%-90%), eliminar ($\leq 70\%$).
- Conciliación: Se detectó una discrepancia significativa con un experto que recomendó eliminar todos los ítems. Se realizaron entrevistas individuales y una reunión de conciliación, lo que permitió depurar, fusionar y reelaborar los ítems, resultando en una versión definitiva de 13 ítems.
- Fase 2 - Validación de Contenido (Criterios de Moriyama): Sobre la versión definitiva, los otros 10 expertos (10 Doctores en Ciencias) evaluaron cada ítem según los cinco criterios de Moriyama: Razonable, Sensible a variaciones, Justificable, Claramente definido y Datos factibles de obtener. La escala de respuesta fue ordinal: Mucho (M), Poco (P), Nada (N).

Para la Fase 2, se realizaron los siguientes análisis cuantitativos:

1. Índice de Validez de Contenido (S-CVI): Se calculó el S-CVI/Ave (Scale-level Content Validity Index, average method) basado en el I-CVI (Ítem-level Content Validity Index). Un I-CVI ≥ 0.78 se consideró aceptable para 10 expertos (Polit et al., 2007).
2. V de Aiken: Se calculó para evaluar la concordancia en la relevancia de los ítems. Un V ≥ 0.70 se consideró aceptable.
3. Kappa de Fleiss (κ): Se calculó para medir la concordancia interjueces más allá del azar, considerando los 65 ítems-criterio (13 ítems \times 5 criterios).

57

Yataria Abreu Vázquez

Ricardo Fernández Rodríguez

René Arenas Gutiérrez

4. Alfa de Cronbach (α): Se calculó como medida de consistencia interna entre las puntuaciones de los expertos, reconociendo sus limitaciones en este contexto específico.

RESULTADOS

Operacionalización y Consenso Inicial

La evaluación inicial de pertinencia mostró que 9 de los 10 expertos recomendaron “Mantener” los 15 ítems originales (90% de acuerdo por ítem). Un experto recomendó “Eliminar” todos los ítems. Tras el proceso de conciliación, que incluyó el análisis cualitativo de las razones detrás de las discrepancias, se refinó la redacción, se fusionaron ítems redundantes y se llegó a una versión definitiva, robusta y consensuada de 13 ítems para la prueba de desempeño y mantener los 3 ítems de la encuesta a directivos. Los que se presentan a continuación.

Prueba de desempeño. (Anexo 1)

Tabla 1.

Prueba de desempeño. Dimensión e indicadores validados

Dimensión	Subdimensión	Indicadores	No de ítem1
Empleo de las TIC en las investigaciones	planificación	Frecuencia con que consulta las fuentes, revisión bibliográfica	53
		Frecuencia con que planifica un recurso	54
	ejecución	Frecuencia con que emplea las TIC en el procesamiento de datos	55
		Frecuencia con que elabora comunicaciones	56
		Frecuencia con que elabora presentaciones	57
		Tipos de recursos empleados en la ejecución del proyecto (ce, WS, IA, RI, O)	58

58

Yataria Abreu Vázquez

Ricardo Fernández Rodríguez

René Arenas Gutiérrez

	Frecuencia con que elabora recursos	59
evaluación	Frecuencia con que emplea las TIC en el procesamiento estadístico	60
	Nivel que alcanza el recurso producido	61
coordinación	Frecuencia de la coordinación interna para el uso de las TIC	62
	Frecuencia de la coordinación externa para el uso de las TIC	63
asesoramiento	Frecuencia con que ofrece asesoramiento	64
	Frecuencia con que recibe asesoramiento	65

Nota: Las escalas de respuesta para frecuencia, nivel y tipos de recursos se detallan en el instrumento original.

Encuesta a directivos validada (Anexo 2)

Validez de Contenido (Criterios de Moriyama)

La evaluación de los 13 ítems definitivos por los otros 10 expertos mostró una predominancia notable de la categoría “Mucho (M)”. Solo se registraron 4 evaluaciones “Poco (P)” en los criterios Razonable y Datos factibles de obtener para los ítems 64 y 65, sobre un total de 650 evaluaciones (13 ítems × 5 criterios × 10 expertos).

Los índices de validez y concordancia calculados fueron:

S-CVI/I-CVI: El I-CVI fue de 1.0 para los ítems 53-63. Los ítems 64 y 65 obtuvieron un I-CVI de 0.9. El S-CVI/Ave resultante fue de 0.983, lo que indica una validez de contenido excelente según los estándares establecidos (5).

V de Aiken: El valor promedio fue de 0.992 (ítems 54-63: V=1.0; ítems 64-65: V=0.95), confirmando una percepción casi unánime de la relevancia de los ítems.

59

Yataria Abreu Vázquez

Ricardo Fernández Rodríguez

René Arenas Gutiérrez



Kappa de Fleiss: El valor calculado fue de $\kappa = 0.825$, lo que indica una concordancia "casi perfecta" entre los evaluadores (Landis & Koch, 1977).

Alfa de Cronbach: El valor fue de $\alpha = 0.72$, sugiriendo una consistencia interna aceptable. Este resultado se interpreta con cautela debido a la naturaleza de los datos de jueces.

En conjunto, estos resultados demuestran que los instrumentos poseen una validez de contenido muy alta y un alto grado de consenso entre el panel de expertos.

DISCUSIÓN

Este estudio logró operacionalizar y validar rigurosamente la variable "Uso de las TIC para la investigación". La estructuración en las cinco dimensiones de la dirección educacional (3) proporciona un marco teórico sólido y pedagógico que trasciende la visión instrumental de las TIC, posicionándolas como elementos estratégicos integrados en la gestión del proceso investigativo.

Los índices de validez de contenido obtenidos ($S-CVI/Ave = 0.983$, V de Aiken = 0.992) son excepcionales y superan los umbrales de excelencia recomendados en la literatura metodológica (Polit et al., 2007). Esto confirma que los 13 ítems de la versión definitiva son percibidos como altamente relevantes, claros y apropiados para medir el constructo. El proceso de conciliación frente a la discrepancia inicial fue crucial para depurar los instrumentos, destacando la importancia de complementar el análisis cuantitativo con la discusión cualitativa entre expertos.

La concordancia interjueces, medida mediante el Kappa de Fleiss ($\kappa = 0.825$), fue "casi perfecta". Este resultado es particularmente robusto, ya que el Kappa de Fleiss es una medida conservadora que corrige el acuerdo esperado por azar. Un

valor tan alto indica un consenso muy fuerte dentro del panel multidisciplinar sobre la validez de los ítems propuestos.

El Alfa de Cronbach ($\alpha = 0.72$), aunque aceptable, debe interpretarse con prudencia. Su cálculo con datos de jueces no es convencional, ya que los evaluadores no son "sujetos" que responden al instrumento, sino jueces que evalúan su calidad. Este valor refleja principalmente la consistencia en las evaluaciones de los expertos, no la consistencia interna del constructo en una población objetivo. Que futuros estudios apliquen estos instrumentos a una muestra de docentes para calcular la confiabilidad de manera apropiada.

CONCLUSIONES

La variable "Uso de las TIC para la investigación" fue operacionalizada exitosamente en cinco dimensiones teóricamente sustentadas y traducida en un conjunto de 12 ítems con una validez de contenido excelente y un alto consenso entre expertos.

La prueba de desempeño y la encuesta a directivos resultantes son instrumentos válidos y confiables para evaluar este constructo en el contexto de la educación médica superior.

Se recomienda que futuras investigaciones apliquen estos instrumentos a una muestra representativa de docentes para realizar análisis psicométricos adicionales, como la validez de constructo mediante análisis factorial confirmatorio y la confiabilidad test-retest, con el fin de consolidar aún más sus propiedades de medición.

El proceso metodológico riguroso descrito sirve como un referente para la validación de instrumentos en el campo de la investigación educativa en salud.

61

Yataria Abreu Vázquez

Ricardo Fernández Rodríguez

René Arenas Gutiérrez



REFERENCIAS BIBLIOGRÁFICAS

- Añorga Morales, J. A. (2014). *La educación avanzada y el mejoramiento profesional: Una teoría para la práctica*. Instituto Superior de Ciencias Médicas de La Habana.
- Barrios Coronel, E., Vargas Viveros, A., Echeverría Cuevas, M., García Pérez, L., & Torales Benítez, J. (2017). Las TIC como herramienta fundamental en la investigación científica. *Revista de Ciencias Sociales, 23*(Especial 2), 1-10.
- Escobar-Pérez, J., & Cuervo-Martínez, Á. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en Medición, 6*, 27-36.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. <https://doi.org/10.2307/2529310>
- Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health, 30*(4), 459–467. <https://doi.org/10.1002/nur.20199>
- Valle Lima, A., Carbonell, M., & Herrera, J. (2015). *Dirección y gestión educativa: Conceptos básicos*. Editorial Pueblo y Educación.

Conflicto de intereses

El autor declara que no existe conflicto de intereses.

62

Yataria Abreu Vázquez

Ricardo Fernández Rodríguez

René Arenas Gutiérrez



Anexo 1. Prueba de desempeño. Dimensión: uso de las TIC en las investigaciones

UNIVERSIDAD DE CIENCIAS MÉDICAS DE LA HABANA
FACULTAD DE CIENCIAS MÉDICAS “JULIO TRIGO LÓPEZ”
PROYECTO: MODELO PARA EL MEJORAMIENTO DEL DESEMPEÑO CIENTÍFICO-
METODOLÓGICO E INVESTIGATIVO DE LOS DOCENTES DE LA FCM “JTL”.
PRUEBA DE DESEMPEÑO Versión 21-4-25

NOMBRES Y APELLIDOS: _____
DPTO. _____ CATEGORÍA DOCENTE: _____
CATEGORÍA INVESTIGATIVA: _____
GRADO ACADÉMICO: _____ GRADO CIENTÍFICO: _____

Estimados colegas un grupo de docentes de la institución, se encuentra realizando una investigación sobre el desempeño investigativo en la Facultad de Ciencias Médicas “Julio Trigo López”, para ello se hace necesario que se autoevalúen en el último año de trabajo (2024), atendiendo a los indicadores propuestos para cada dimensión del trabajo, estas son: planificación, ejecución, evaluación, coordinación y asesoramiento, esperamos su honesta contribución al desarrollo de la investigación.

Para la evaluación de las dimensiones se establecen las categorías siguientes:

frecuencia

1 Nunca, 2 Pocas veces, 3 Algunas veces, 4 Casi siempre, 5 Siempre

Nivel

1 Ninguno, 2 Poco elevado, 3 medio, 4 elevado, 5 Muy elevado

Temáticas:

Muchas gracias por su colaboración.

Recuerde se autoevalúa el último año (2024). Para la evaluación de las dimensiones se establecen las categorías siguientes:

frecuencia

1 Nunca, 2 Pocas veces, 3 Algunas veces, 4 Casi siempre, 5 Siempre

Nivel

1 Ninguno, 2 Poco elevado, 3 medio, 4 elevado, 5 Muy elevado

Tipos de recursos

1 correo electrónico (**ce**), 2 whatsapp (**WS**), 3 inteligencia artificial (**IA**), 4 redes de investigadores (**RI**), 5 otros (**O**)

Dimensión: Empleo de las TIC en las investigaciones

No	Subdimensión	Indicadores	1	2	3	4	5
53	planificación	Frecuencia con que consulta las fuentes, revisión bibliográfica					
54		Frecuencia con que planifica un recurso					
55	ejecución	Frecuencia con que emplea las TIC en el procesamiento de datos					
56		Frecuencia con que elabora comunicaciones					

63

Yataria Abreu Vázquez

Ricardo Fernández Rodríguez

René Arenas Gutiérrez



57		Frecuencia con que elabora presentaciones
58		Tipos de recursos empleados en la ejecución del proyecto (ce, WS, IA, RI, O)
59		Frecuencia con que elabora recursos
60	evaluación	Frecuencia con que emplea las TIC en el procesamiento estadístico
61		Nivel que alcanza el recurso producido
62	coordinación	Frecuencia de la coordinación interna para el uso de las TIC
63		Frecuencia de la coordinación externa para el uso de las TIC
64	asesoramiento	Frecuencia con que ofrece asesoramiento
65		Frecuencia con que recibe asesoramiento

Anexo 2. Encuesta a directivos

**UNIVERSIDAD DE CIENCIAS MÉDICAS DE LA HABANA
FACULTAD DE CIENCIAS MÉDICAS “JULIO TRIGO LÓPEZ”
PROYECTO: MODELO PARA EL MEJORAMIENTO DEL DESEMPEÑO CIENTÍFICO-
METODOLÓGICO E INVESTIGATIVO DE LOS DOCENTES DE LA FCM “JTL”.
Encuesta a directivos**

DPTO. _____ CATEGORÍA DOCENTE: _____

CATEGORÍA INVESTIGATIVA: _____

GRADO ACADÉMICO: _____ GRADO CIENTÍFICO: _____

Estimados colegas un grupo de docentes de la institución, se encuentra realizando una investigación sobre el desempeño investigativo en la Facultad de Ciencias Médicas “Julio Trigo López”, para ello se hace necesario emitir sus criterios atendiendo a los indicadores propuestos para cada dimensión del trabajo, estas son: planificación, ejecución, evaluación, coordinación y asesoramiento, esperamos su honesta contribución al desarrollo de la investigación. Aclaramos que en la primera pregunta de cada dimensión solo debe marcarse una opción en cada etapa.

Dimensión: Empleo de las TIC en las investigaciones

1. ¿Cómo evalúa el empleo de las TIC en las investigaciones que se desarrollan en su área de trabajo?:

La planificación: ___ B ___ R ___ M

La ejecución ___ B ___ R ___ M

La evaluación ___ B ___ R ___ M

La coordinación ___ B ___ R ___ M

El asesoramiento ___ B ___ R ___ M

2. ¿Qué aspectos señalaría como potencialidades y cuáles como falencias?

Potencialidades: _____

Falencias: _____

3. ¿Qué sugerencias puede ofrecer para obtener mejores resultados en este proceso?
