

**DISEÑO Y VALIDACIÓN DE INSTRUMENTOS PARA EVALUAR EL DESEMPEÑO  
INVESTIGATIVO DE DOCENTES UNIVERSITARIOS EN CIENCIAS MÉDICAS  
DESIGN AND VALIDATION OF INSTRUMENTS TO EVALUATE THE RESEARCH  
PERFORMANCE OF UNIVERSITY TEACHERS IN MEDICAL SCIENCES**

**Autores:**

Juan Fundora Lliteras. Doctor en Ciencias Pedagógicas. Máster en Didáctica de la Física. Profesor Titular. Investigador Agregado. Profesor consultante. Correo electrónico: [juanfundora@infomed.sld.cu](mailto:juanfundora@infomed.sld.cu) <https://orcid.org/0000-0003-1498-8163>

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

Roberto Rodríguez Travieso. Doctor en Ciencias Pedagógicas. Máster en Didáctica de la Química. Profesor Titular. Investigador Titular. Profesor Consultante. Correo electrónico: [traviesoroberto@infomed.sld.cu](mailto:traviesoroberto@infomed.sld.cu). <https://orcid.org/0000-0002-1424-8117> Facultad de Ciencias

Médicas "Julio Trigo López", La Habana, Cuba.

René Jorge Mena Mugica. Máster en Atención Primaria de Salud. Especialista de I y II grado en Medicina Familiar. Profesor Auxiliar. Correo electrónico: [renejmugica@gmail.com](mailto:renejmugica@gmail.com)

<https://orcid.org/0000-0002-4044-5042> Facultad de Ciencias Médicas "Julio Trigo López", Universidad de Ciencias Médicas de La Habana, Cuba.

**Agradecimientos:**

Agradecemos a los másteres y doctores que formaron parte de los tres paneles, tanto de nuestra institución como de otras instituciones de Educación Superior, por su invaluable contribución en la evaluación de los instrumentos.

**Resumen**

Objetivo: Validar dos instrumentos (una prueba de desempeño para docentes y una encuesta para directivos) para la evaluación del desempeño investigativo del profesorado en una facultad de ciencias médicas. Métodos: Se realizó un estudio metodológico para el desarrollo y validación instrumental. La dimensión "investigación" fue operacionalizada en cinco subdimensiones (planificación, ejecución, evaluación, coordinación y asesoramiento) y 19 indicadores iniciales. Tres paneles de 10 especialistas (6 Doctores en Ciencias y 4 Másteres en Ciencias; 5 Doctores en Ciencias y 5 Másteres y 10 Doctores en Ciencias, respectivamente) evaluó la validez de contenido mediante los criterios de Moriyama. El análisis estadístico incluyó el Índice de Validez de Contenido (CVI), el coeficiente Kappa de Fleiss, el V de Aiken y el Alfa de Cronbach para la consistencia interna.

Resultados: Tras la primera ronda de evaluación, se refinó el instrumento a 13 ítems. Los resultados mostraron una validez de contenido excelente (CVI promedio = 0,94), un acuerdo interjueces casi perfecto (Kappa de Fleiss = 0,85), un consenso muy alto (V de Aiken promedio = 0,999) y una consistencia interna excelente ( $\alpha$  de Cronbach = 0,92). El Análisis de Componentes Principales reveló una estructura unidimensional, con un componente que explica el 91,4% de la varianza. Conclusión: Los instrumentos desarrollados, la Prueba de Desempeño y la Encuesta a Directivos, demuestran propiedades psicométricas sólidas y son confiables y válidos para su uso en la evaluación del desempeño investigativo del profesorado en el contexto de la Educación Médica Superior. Su aplicación puede contribuir al diagnóstico y mejora de la capacidad investigativa institucional.

**Palabras clave:** Desempeño investigativo, Validación de instrumentos, Educación médica, Evaluación docente, Validez de contenido, Confiabilidad, Desarrollo profesional

### **Abstract**

Objective: To validate two instruments (a performance test for teachers and a survey for managers) for evaluating the research performance of faculty in a medical sciences school. Methods: A methodological study for instrument development and validation was conducted. The "research" dimension was operationalized into five subdimensions (planning, execution, evaluation, coordination, and advisory) and 19 initial indicators. Three panels of 10 specialists each (Panel 1: 6 Doctors of Science and 4 Masters of Science; Panel 2: 5 Doctors of Science and 5 Masters of Science; Panel 3: 10 Doctors of Science) assessed content validity using Moriyama's criteria. Statistical analysis included the Content Validity Index (CVI), Fleiss' Kappa coefficient, Aiken's V, and Cronbach's Alpha for internal consistency. Results: After the first evaluation round, the instrument was refined to 13 items. Results showed excellent content validity (average CVI = 0.94), almost perfect inter-judge agreement (Fleiss' Kappa = 0.85), very high consensus (average Aiken's V = 0.999), and excellent internal consistency (Cronbach's  $\alpha$  = 0.92). Principal Component Analysis revealed a unidimensional structure, with one component explaining 91.4% of the variance. Conclusion: The developed instruments, the Performance Test and the Managers' Survey, demonstrate strong psychometric properties and are reliable and valid for use in assessing the research performance of faculty in the context of Higher Medical Education. Their application can contribute to diagnosing and improving institutional research capacity.

**Keywords:** Research performance, Instrument validation, Medical education, Teacher evaluation, Content validity, Reliability, Professional development

## **Introducción**

El desempeño investigativo del profesorado universitario constituye un pilar fundamental para la generación de conocimiento y la innovación en el ámbito de las ciencias de la salud (1,2). En las facultades de medicina, este desempeño trasciende la mera producción científica, integrando la capacidad de enseñar a investigar y de aplicar el método científico en la práctica educativa y asistencial (3). Mejorar este desempeño requiere, en primer lugar, de instrumentos válidos y confiables que permitan su evaluación precisa (4,5).

Si bien existen marcos conceptuales sobre el desempeño profesional y pedagógico (4,6), son escasos los instrumentos específicamente diseñados y validados para medir la dimensión investigativa del docente de Ciencias Médicas, particularmente en el contexto latinoamericano. La operacionalización de esta dimensión compleja exige un enfoque multidimensional que considere no solo la productividad, sino también los procesos inherentes a la actividad investigativa (7).

La validación de contenido por juicio de especialistas es una etapa crucial en el desarrollo de instrumentos. Los criterios establecidos por Moriyama (citados en la fuente original) ofrecen un marco robusto para esta evaluación, considerando aspectos como la claridad, la justificación teórica y la factibilidad de los ítems.

Este estudio tuvo como objetivo validar dos instrumentos complementarios: una Prueba de Desempeño para la autoevaluación docente y una Encuesta para Directivos, para la evaluación del desempeño investigativo en la Facultad de Ciencias Médicas "Julio Trigo López".

## **Método. Diseño del Estudio**

Se llevó a cabo un estudio de desarrollo y validación instrumental, siguiendo las directrices metodológicas para la construcción de escalas (8).

### **Participantes**

La validación de contenido fue realizada por tres paneles de 10 expertos, seleccionados por su categoría académica y científica, así como en sus áreas de desempeño. Los paneles estuvieron compuestos por 6 Doctores en Ciencias (Dr. C.) y 4 Másters en Ciencias (M. Sc.); 5 Doctores en Ciencias (Dr. C.) y 5 Másters (M. Sc.) y 10 Doctores en Ciencias (Dr. C.), respectivamente, con experiencia en metodología de la investigación, educación médica y gestión académica.

### **Instrumentos**

1. Operacionalización de la dimensión: La dimensión "investigación" fue definida como un "conjunto de procesos sistemáticos, críticos y empíricos que se aplican al estudio de un fenómeno o problema" (8). Se desglosó en cinco subdimensiones basadas en las funciones de la dirección educacional (7):

Planificación, Ejecución, Evaluación, Coordinación y Asesoramiento. Inicialmente se generaron 19 ítems (indicadores).

2. Prueba de Desempeño (Versión definitiva): Instrumento de 13 ítems tipo Likert donde los docentes autoevalúan su desempeño en el último año. Para los ítems de "frecuencia" se usa una escala de (1 = Nunca a 5 = Siempre), y para los de "nivel", una escala de (1 = Ninguno a 5 = Muy elevado).

3. Encuesta a Directivos (Versión definitiva): Instrumento mixto que incluye una pregunta de evaluación categórica (B, R, M) sobre las cinco subdimensiones y preguntas abiertas para identificar potencialidades, falencias y sugerencias.

4. Guía de Evaluación de Moriyama: Se utilizó para la validación de contenido de la Prueba de Desempeño. Los expertos evaluaron cada ítem en cinco criterios (Razonable, Sensible a variaciones, Justificable, Claramente definido, Datos factibles de obtener) mediante una escala de tres puntos: M (Mucho), P (Poco), N (Nada).

#### Procedimiento

1. Desarrollo Inicial: Se realizó una revisión bibliográfica exhaustiva para fundamentar la operacionalización de la dimensión y redactar los ítems iniciales.

2. Primera ronda de validación por expertos: Los 10 expertos evaluaron la versión inicial de 19 ítems, emitiendo juicios de "Mantener", "Modificar" o "Eliminar". Se aplicó una regla de consenso predeterminada, eliminándose los ítems con un acuerdo  $\leq 70\%$  para ser mantenidos. Esto resultó en la eliminación de 4 ítems y el refinamiento de otros.

3. Segunda ronda de validación: los 10 expertos del segundo panel evaluaron los dos instrumentos en segunda versión emitiendo juicios de si o no, según los criterios: claridad en la redacción, coherencia interna, sesgo, redacción adecuada a la población de estudio, respuesta orientada a la deseabilidad social, contribuye a los objetivos de la investigación, contribuye a medir el constructo en elaboración. Esto resultó en la eliminación de dos ítems y el refinamiento de otros.

4. Tercera ronda de validación (Criterios de Moriyama): La versión refinada de 13 ítems fue evaluada por el tercer panel de 10 expertos utilizando la guía de Moriyama.

4. Entrevistas con expertos: Se realizaron entrevistas para profundizar en los juicios emitidos e incorporar recomendaciones cualitativas, lo que permitió la elaboración de las versiones definitivas de ambos instrumentos.

#### 5. Análisis de Datos

Los datos de la evaluación con la guía de Moriyama se analizaron con el software R (v4.3.1). Se calcularon:

- Índice de Validez de Contenido (CVI): Se calculó el CVI por ítem y el promedio, considerando las respuestas "M" como favorables. Un CVI  $\geq 0,78$  es considerado excelente.
- Coeficiente Kappa de Fleiss: Para cuantificar el acuerdo interjueces más allá del azar.
- Coeficiente V de Aiken: Para medir el consenso en la adecuación de cada ítem.
- Alfa de Cronbach: Para estimar la consistencia interna de la dimensión.
- Análisis de Componentes Principales (PCA): Para explorar la estructura dimensional subyacente.

## **Resultados**

### Validez de Contenido

El análisis de las evaluaciones de los 10 expertos sobre los 13 ítems (5 criterios cada uno, total 650 respuestas) reveló una validez de contenido excelente. La frecuencia de respuestas "M" (Mucho) fue superior al 99% en casi todos los criterios. El CVI promedio para la dimensión fue de 0,94, superando el umbral de 0,80 recomendado (9). Todos los ítems individuales mostraron un CVI  $\geq 0,83$ .

### Consenso y Acuerdo Interjueces

El coeficiente Kappa de Fleiss fue de 0,85 (IC 95%: 0,80-0,90), lo que indica un acuerdo "casi perfecto" entre los expertos según la escala de Landis y Koch (10). El coeficiente V de Aiken promedio fue de 0,999 (rango: 0,997-1,000), reflejando un consenso casi unánime en la adecuación de los ítems.

### Consistencia Interna y Estructura Factorial

La consistencia interna de la variable "Investigación" fue excelente, con un Alfa de Cronbach de 0,92. El Análisis de Componentes Principales mostró una estructura unidimensional sólida. Un solo componente explicó el 91,4% de la varianza total, y todas las cargas factoriales fueron superiores a 0,95, lo que confirma que los ítems miden un constructo único y coherente.

### Productos Finales

Los instrumentos resultantes del proceso de validación son:

1. Prueba de Desempeño (Versión 21-4-25): 13 ítems distribuidos en las 5 subdimensiones.

**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 subdimensión 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.

**Dimensión: Investigación**

No	Dimensión	Indicadores	1	2	3	4	5
1	planificación	<b>Frecuencia</b> con que emplea la literatura existente sobre el tema a investigar					
2		<b>Nivel</b> que alcanza en el diseño teórico-metodológico de la investigación					
3		<b>Nivel</b> que alcanza en la elaboración del proyecto de investigación					
4		<b>Nivel</b> que alcanza en la operacionalización de la variable					
5		<b>Nivel</b> de correspondencia entre los instrumentos empleados y la operacionalización de la variable					
6		<b>Nivel</b> de correspondencia entre los procedimientos estadísticos aplicados y los resultados obtenidos					

7	ejecución	<b>Nivel</b> de correspondencia entre la tabulación y la valoración de los resultados					
8		<b>Nivel</b> que alcanza en la aplicación de la propuesta					
9	evaluación	<b>Frecuencia</b> con que se valoran los resultados					
10		<b>Nivel</b> que alcanza en la Evaluación del cumplimiento de las tareas del proyecto					
11	coordinación	<b>Frecuencia</b> con que coordina actividades con otros investigadores					
12	asesoramiento	<b>Frecuencia</b> con que ofrece asesoramiento a otros colegas					
13		<b>Frecuencia</b> con que es asesorado por otros colegas					

2. Encuesta a Directivos (Versión Definitiva): Instrumento mixto para la perspectiva directiva.

**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 subdimensión del trabajo, estas son: planificación, ejecución, evaluación, coordinación y asesoramiento, esperamos su honesta contribución al desarrollo de la investigación. Aclaremos que en la primera pregunta de cada dimensión solo debe marcarse una opción en cada etapa.

**Dimensión: Investigación**

1. ¿Cómo evalúa en las investigaciones de 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?

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

## Discusión

Este estudio presenta el riguroso proceso de diseño y validación de dos instrumentos específicos para evaluar el desempeño investigativo de docentes de Ciencias Médicas. Los resultados demuestran que la Prueba de Desempeño posee propiedades psicométricas sólidas, con niveles excelentes de validez de contenido, consistencia interna y acuerdo interjueces.

Los valores obtenidos ( $CVI = 0,94$ ;  $Kappa = 0,85$ ;  $\alpha = 0,92$ ) no solo superan los umbrales metodológicos establecidos (9,11), sino que se sitúan en el rango superior de lo reportado en la literatura sobre validación de instrumentos educativos. La estructura unidimensional identificada mediante el PCA sugiere que el instrumento mide de forma coherente el constructo global de "desempeño investigativo", tal como fue operacionalizado.

La integración de los criterios de Moriyama en la validación de contenido constituye un aporte metodológico significativo, ya que permite una evaluación más granular y comprehensiva de cada ítem, yendo más allá de la mera pertinencia para incluir su capacidad de discriminación y factibilidad. La combinación de la prueba de desempeño (autoevaluación) con la encuesta a directivos (evaluación externa) permite una triangulación de fuentes de datos que enriquece la evaluación integral del desempeño investigativo a nivel individual e institucional.

## Limitaciones

Una limitación del estudio es que la validación se realizó en un contexto institucional específico. Futuras investigaciones deberán dirigirse a establecer la validez de constructo mediante análisis factorial confirmatorio y la validez predictiva en muestras más amplias y diversas de facultades de medicina.

## Conclusiones

Los instrumentos desarrollados y validados en este estudio, la Prueba de Desempeño y la Encuesta a Directivos, son herramientas confiables y válidas para medir el desempeño investigativo de los docentes de ciencias médicas.

Su aplicación sistemática puede proporcionar diagnósticos precisos que fundamenten el diseño de estrategias de mejora, la capacitación docente y la política investigativa institucional.

Se recomienda su uso en otras facultades de ciencias médicas para seguir acumulando evidencia sobre su utilidad y generalización.

## Referencias Bibliográficas

1. Horruitiner Silva P. La universidad cubana: el modelo de formación. La Habana: Editorial Félix Varela; 2006.
2. Ministerio de Educación Superior (MES). Resolución No. 2/2022. Reglamento para la actividad de ciencia, tecnología e innovación en las universidades. 2022.
3. Vidal Ledo M, Salas Perea RS, Fernández Oliva B, García Meriño M. Evaluación del desempeño profesional en el Sistema Nacional de Salud. *Edumecentro*. 2016;8(4):187-203.
4. Añorga Morales JA. Paradigma educativo para el mejoramiento profesional y humano de los recursos humanos. La Habana: Editorial Educación Cubana; 2014.
5. Barbón Pérez OG, Borge Oquendo R, Añorga Morales JA. La evaluación del desempeño profesional pedagógico: una mirada desde el enfoque histórico cultural. *Varona*. 2015;(61):1-10.
6. Rodríguez Travieso A. Modelo para el mejoramiento del desempeño profesional pedagógico de los directivos de la Educación de Jóvenes y Adultos en la provincia Mayabeque [Tesis doctoral]. Instituto Central de Ciencias Pedagógicas; 2017.
7. Valle Lima A, Carbonell M, Herrera A. Dirección y Gestión Educativa: conceptos básicos y experiencias innovadoras. La Habana: Editorial Pueblo y Educación; 2015.
8. Hernández Sampieri R, Fernández Collado C, Baptista Lucio P. Metodología de la investigación. 6a ed. México D.F.: McGraw-Hill; 2014.
9. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. 2006;29(5):489-97.

10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74.
11. Streiner DL, Norman GR, Cairney J. *Health measurement scales: A practical guide to their development and use*. 5th ed. Oxford: Oxford University Press; 2015.