

MARCO PARA LA ORIENTACIÓN Y EVALUACIÓN DEL RIGOR CIENTÍFICO EN TESIS DE MAESTRÍA EN EDUCACIÓN

Framework for the Orientation and Evaluation of Scientific Rigor in Master's Thesis Research in Education

Ramón Alberto López Ynoa

Universidad Internacional Iberoamericana (ramon.lopez@doctorado.unib.org) (<https://orcid.org/0009-0002-9627-493x>)

Óscar Casanova López

Universidad de Zaragoza (ocasanov@unizar.es) (<https://orcid.org/0000-0002-8263-3447>)

información del manuscrito:

Recibido/Received: 18/11/25

Revisado/Reviewed: 03/02/26

Aceptado/Accepted: 22/06/26

RESUMEN

Palabras clave:

Rigor científico
Tesis de maestría
Investigación educativa
Validación de instrumentos
Propiedades psicométricas

El propósito de esta investigación fue proponer un marco teórico-metodológico con criterios e indicadores válidos y confiables para orientar y evaluar el rigor científico de las tesis de maestría en educación de la Universidad Autónoma de Santo Domingo, Recinto Barahona. Se aplicó una metodología postpositivista sustentada en el realismo crítico y la epistemología objetivista, combinando análisis cuantitativos y cualitativos. La construcción del Instrumento de Evaluación del Rigor Científico de las Tesis de Maestría en Educación (IERCITEM) incluyó juicio de expertos, validez de contenido, aplicación piloto para la fiabilidad y validación criterial. Participaron siete jueces y se analizó una muestra de 53 tesis correspondientes al período 2015-2021. Los resultados evidenciaron una validez de contenido sobresaliente (CVR = 0.95; I-CVI = 0.98; V de Aiken = 0.94; ICE = 0.95), alta fiabilidad (α = 0.91; ICC = 0.87; G = 0.85) y una correspondencia significativa con el referente utilizado (r = 0.86; κ = 0.72; ICC = 0.89). Asimismo, se confirmaron seis criterios interdependientes: metodológico, axiológico, praxiológico, epistemológico, ontológico y formal. Estos hallazgos respaldan una concepción multidimensional del rigor científico, en la que la calidad de la investigación depende de la coherencia entre fundamentos teóricos, éticos, metodológicos y formales, más allá del cumplimiento técnico de procedimientos. Aunque la validación se circunscribió a una institución y área disciplinar específicas, la evidencia obtenida sugiere que el IERCITEM constituye una herramienta sólida, válida y replicable para fortalecer la cultura investigativa y los procesos de aseguramiento de la calidad académica en el posgrado.

ABSTRACT

Keywords:

Scientific rigor

The purpose of this study was to propose a theoretical-methodological framework comprising valid and reliable criteria and indicators to guide and evaluate the scientific rigor of master's theses in education at the

Master's theses Educational research Instrument validation Psychometric properties	Autonomous University of Santo Domingo, Barahona Campus. A post-positivist methodology grounded in critical realism and objectivist epistemology was employed, integrating quantitative and qualitative analyses. The development of the Instrument for Evaluating Scientific Rigor in Master's Theses in Education (IERCITEM) involved expert judgment, content validation, pilot testing for reliability, and criterion-related validation. Seven expert judges participated, and a sample of 53 theses produced between 2015 and 2021 was analyzed. The results demonstrated outstanding content validity (CVR = 0.95; I-CVI = 0.98; Aiken's V = 0.94; ICE = 0.95), high reliability (α = 0.91; ICC = 0.87; G = 0.85), and significant correspondence with the reference instrument (r = 0.86; κ = 0.72; ICC = 0.89). Furthermore, six interdependent dimensions were confirmed: methodological, axiological, praxiological, epistemological, ontological, and formal. These findings support a multidimensional conception of scientific rigor in which research quality depends on the coherence among theoretical, ethical, methodological, and formal foundations, extending beyond the mere technical application of procedures. Although the validation was limited to a single institution and disciplinary field, the evidence suggests that IERCITEM is a robust, valid, and replicable instrument capable of strengthening research culture and enhancing academic quality assurance processes in graduate education.
---	--

Introducción

La preocupación por la calidad de las investigaciones educativas ha ido cobrando creciente relevancia en los últimos años debido, entre otras consideraciones, a su escasa reproducibilidad y a la reducida generación de conocimiento original en este campo. Así, las investigaciones realizadas sobre esta temática han puesto de manifiesto las debilidades estructurales de la producción académica, relacionadas con su insuficiente rigor científico y metodológico. En el caso de la República Dominicana, esta situación ha limitado la mejora de la práctica educativa y la respuesta académica a las demandas sociales en el ámbito educativo.

Por calidad científica se entiende a que un estudio aporte originalidad, relevancia y contribución al estado del arte, evaluado con criterios y estándares como visibilidad, utilidad, rigor metodológico y transparencia (Arguimbau, 2023; Correa Hincapié et al., 2023). El rigor científico, por su parte, implica el apego estricto a un conjunto de normas y procedimientos que aseguran resultados válidos, confiables y replicables (Guillén Díaz y Sanz Trigueros, 2021).

Los estudios realizados hasta la fecha sobre el rigor científico de las tesis de maestría en educación han analizado aspectos y elementos estructurales específicos, pero pocos han examinado las fundamentaciones teóricas, la calidad y el rigor desde una perspectiva global (Díaz Torres, 2024; Ge et al., 2024). Se constata deficiencias en la calidad, en el diseño teórico-metodológico, el análisis del estado del arte y en la estructuración del aparato crítico en la fase de elaboración del proyecto, en la prueba de hipótesis, la discusión y en el análisis estadístico (Francia y Figueroa Morán, 2022; Francia Hernández, 2023; Huayta Meza, 2021).

También, se han detectado deficiencias en el manejo de los fundamentos teórico-metodológicos de la investigación, incongruencias entre los antecedentes, los referentes teóricos y el planteamiento del problema, así como dificultades para constatar la novedad científica (González Martínez et al., 2022; Nina Condori, 2022).

Las investigaciones en este campo señalan las deficiencias de rigor científico de las tesis de maestría en educación y las vinculan a la ausencia de un sistema de criterios e indicadores de referencia para desarrollarlas y evaluarlas. Se destaca la falta de validación de los instrumentos utilizados en los estudios que los presentan. No obstante, la mayoría de los estudios no ofrecen estándares, criterios ni indicadores de referencia al abordar el estudio de un objeto de investigación educativa.

En respuesta, se han realizado aportaciones teóricas y se han propuesto instrumentos concebidos para evaluar el informe de la tesis de pregrado y posgrado (Aliaga-Pacora et al., 2023; Ceballos-Almeraya y Tobón, 2019; Vásquez-Antonio et al., 2018). De igual manera, se han propuesto instrumentos que, de algún modo, se vinculan a la estructura de una tesis de maestría en educación. Sin embargo, estos instrumentos no están concebidos para servir de guía durante el desarrollo del proceso, sino para evaluar las competencias investigativas del sustentante (Aliaga-Pacora et al., 2023; Cabrera et al., 2023).

Otras propuestas (Chambi Mosco, 2017; Soto, 2021) que plantean soluciones a la situación descrita no pueden considerarse marcos integrales de valoración del rigor científico, ya que limitan sus aportaciones al aspecto metodológico y minusvaloran elementos que deberían tener mayor peso específico. En los estudios persiste el hecho de que los aspectos éticos de la investigación, por ejemplo, no se consideran parte de los criterios de rigor científico (Ochoa Sierra y Moreno Mosquera, 2019).

De esta manera, las grillas que se proponen en estudios que van más allá de las recomendaciones retóricas sobre el diseño de instrumentos, se consideran insuficientes. Esto es así porque no están diseñadas en respuesta a un esquema marco en el que se entrelacen sistémicamente los criterios que definen la calidad y el rigor científico de una investigación académica.

Estos instrumentos, reflejan, en consecuencia, una lógica que fragmenta la integralidad del hecho científico y presentan un desequilibrio a favor de criterios tradicionales de la actividad investigativa académica. Entre las recomendaciones de instrumentos para evaluar el rigor de las tesis, se encuentran cuestionarios y rúbricas analíticas socioformativas (Cardoso y Cerecedo, 2019; George-Reyes y Ramírez-Martinell, 2019; Soto, 2021; Tapia et al., 2018).

En la República Dominicana prevalecen las normativas institucionales con una débil presencia de estudios empíricos con instrumentos validados y análisis crítico. Esta situación produce una brecha que impide comprobar la satisfacción de las reglamentaciones en las principales universidades dominicanas creando un vacío investigativo en educación.

Es evidente la ausencia de propuestas transferibles de instrumentos que sirvan de guía para el desarrollo y evaluación de las investigaciones académicas en el nivel de maestría en educación, y que garanticen la calidad y el rigor científico de las mismas, y que, por ende, avalen la calidad académica institucional.

La falta de orientación para dotar de rigor científico a las tesis de maestría en educación durante su desarrollo, justifica la construcción de un marco válido y confiable de indicadores de calidad que garantice la cientificidad de dichas tesis. Este marco debe garantizar la calidad académica de los programas de maestría y la calidad científica de los trabajos de finalización de maestría, particularmente, en nuestro país, que evidencia una inexplicable ausencia en esta línea de investigación.

Por ello, el objetivo de esta investigación es proponer un sistema de criterios con indicadores válidos y confiables que sirva de marco de aseguramiento de la calidad de las tesis de maestrías en educación en la Universidad Autónoma de Santo Domingo Recinto Barahona, cumpliendo con altos estándares de rigor científico. Al proponer un referente teórico-metodológico de adecuación al rigor científico, se busca aumentar la credibilidad de estos estudios y garantizar la reproducibilidad de sus resultados.

Desde una perspectiva teórico-conceptual, el diseño de este sistema de calidad se basa en las aportaciones de Guillén Díaz y Sanz Trigueros (2021) y de Deroncele Acosta et al. (2021). Guillén Díaz y Sanz Trigueros (2021) proponen una visión estructurada de la calidad investigativa anclada en tres dimensiones: calidad de la investigación, calidad en la investigación y la excelencia.

Para estos autores el rigor científico está vinculado a la calidad de la investigación, determinada por la solidez de sus conclusiones y su aportación a nuevos conocimientos, o al incremento de los ya existentes. Asimismo, enfocan la cientificidad desde la calidad, combinando aspectos metodológicos, relevancia y repercusión práctica de la investigación.

Deroncele Acosta et al. (2021), por su parte, aportan una propuesta de competencia epistémica que integra cinco criterios claves para evaluar el rigor científico: metodológico, axiológico, praxiológico, epistemológico y ontológico. Como novedad, en este estudio se ha vinculado *aspectos formales* como componente diferenciado, con su peso específico. De esta manera, resulta un sistema integral de seis macro criterios orientado al acompañamiento y a la evaluación que promueve la calidad científica de las tesis en el nivel de maestría.

El esquema contribuye a la comprensión teórica del rigor científico y está concebido para que sea referente transferible en la práctica investigativa y evaluativa en el nivel académico de postgrado. El peso específico de cada criterio viene dado por la carga de ítems con la que cuenta el instrumento diseñado para la operatividad del sistema, lo que dota a la propuesta de transparencia metodológica y fortalece su aplicabilidad práctica.

En este contexto, se presenta el *Instrumento de Evaluación del Rigor Científico de Tesis de Maestría en Educación* (IERCITEM), un instrumento cuantitativo-descriptivo que asume los cinco criterios teórico-metodológicos de Deroncele Acosta et al. (2021) y se le integra Aspectos Formales, considerados fundamentales para definir el rigor científico de una tesis de maestría en educación. Se concibe como una herramienta de evaluación rigurosa, con un enfoque integral sustentado teóricamente, que contribuye a la mejora continua del proceso de investigación y referente del informe final.

Método

Naturaleza teórica y ruta metodológica en la construcción del IERCITEM

El IERCITEM fue concebido desde el paradigma postpositivista empírico-analítico, integrando aportaciones del realismo crítico y la epistemología objetivista. Esta base permitió desarrollar un instrumento teórica y metodológicamente sólido, con criterios e indicadores que orientan la producción científica.

El realismo crítico proporcionó el marco epistemológico para analizar el rigor científico de las tesis, ya que reconoce la existencia de estándares de calidad. Presenta evidencia cuando se analizan las tesis para determinar en qué medida cumplen con estos criterios, para lo que utiliza métodos de análisis cuantitativos y cualitativos. El objetivismo epistemológico aseguró la coherencia interna del instrumento, su validez y la replicabilidad de los resultados, consolidando un sistema referencial que combina análisis cuantitativo y cualitativo.

Ruta metodológica

El proceso de elaboración del IERCITEM siguió una ruta metodológica estructurada en etapas secuenciadas: etapa previa: versión inicial del instrumento; etapa 1: intervención de los jueces; etapa 2: intervención de los evaluadores; y etapa 3: post-validación final: validez de constructo y criterial.

En la etapa previa, se definió con claridad el constructo central y las variables a evaluar, precisándose las dimensiones, subdimensiones e indicadores observables que sirvieron de base para el diseño inicial del instrumento. El instrumento se operacionalizó con un formato de respuesta tipo Likert de cinco niveles ponderables de valoración. En esta etapa se construyó una versión preliminar del instrumento.

En la etapa 1 participaron siete jueces especializados en educación y metodología de la investigación. Se garantizó la validez de contenido evaluando cada ítem en los aspectos de pertinencia, claridad y suficiencia. Se identificaron fortalezas y debilidades en la redacción de los ítems, dando lugar a un proceso de revisión en el que se ajustaron los ítems conforme a la retroalimentación recibida. El resultado fue una versión del instrumento lista para someterse a una prueba empírica.

A continuación, la etapa 2 estuvo orientada a la evaluación psicométrica del instrumento. Para ello, se seleccionó una muestra de 53 tesis de un total de 106 tesis aprobadas en la UASD Recinto Barahona durante el período 2015-2021. La selección se efectuó mediante muestreo probabilístico proporcional por año, con selección aleatoria dentro de cada estrato. Participaron siete evaluadores, quienes recibieron capacitación previa para garantizar la aplicación estandarizada del instrumento. Posteriormente, se realizó la prueba piloto, lo que permitió la recolección de datos empíricos en condiciones reales.

Con los datos obtenidos se evaluaron la fiabilidad, la validez de constructo y la validez de criterio del instrumento. La validez de constructo se examinó mediante análisis factorial, previa comprobación de los supuestos de adecuación muestral a través del índice Kaiser-Meyer-Olkin (KMO) y de la prueba de esfericidad de Bartlett, cuyos resultados evidenciaron la idoneidad de los datos para este tipo de análisis.

Finalmente, en la etapa 3 se realizaron los ajustes derivados de los análisis psicométricos y de la revisión integral de los resultados obtenidos. Esta fase permitió consolidar la versión definitiva del IERCITEM, validado para su aplicación como referente en el desarrollo de investigaciones y en la evaluación del rigor científico de tesis de maestría en educación.

Una vez concluida esta etapa, se preparó un manual de instrucciones, de aplicación y de calificación, con el fin de garantizar la estandarización en la aplicación del instrumento en diferentes contextos.

Descripción del IERCITEM y proceso de validación

En la construcción del instrumento se tuvieron en cuenta seis criterios: Metodológico, Axiológico, Praxiológico, Epistemológico, Ontológico (Deroncele Acosta et al., 2021), y Aspectos Formales. Se diseñó tras una profunda revisión bibliográfica y el análisis de contribuciones de autores que han abordado aspectos similares en el ámbito educativo.

El instrumento está diseñado con una escala tipo Likert, con cinco niveles de valoración ascendentes, presenta equivalencias porcentuales, lo que permite realizar de manera precisa análisis estadísticos comparativos inter ítems e inter criterios. La presencia de subítems refuerza la objetividad en la calificación, ya que cada criterio e indicador recibe una valoración cuantitativa-cualitativa relativa según el número de subítems que lo componen.

La Tabla 1 describe el peso específico de cada criterio, determinado por la cantidad de ítems, su carga porcentual y su valoración ponderada de rigor científico en el IERCITEM. La valoración de cada criterio, al igual que la tesis en general, se detalla en una escala valorativa de 1-5, con una puntuación/100.0 y su equivalencia porcentual (%).

La ponderación cualitativa de cada criterio es: Muy alto (5 puntos = 96.0-100.0%), Alto (4 puntos = 80.0-95.0%), Medio (3 puntos = 70.0-79.0%), Bajo (2 puntos $\geq$ 69.0%), y Muy bajo/Inaceptable (1 punto $\geq$ 19.0%). La calificación general de la tesis viene dada por la media de las puntuaciones de cada uno de sus criterios.

Tabla 1*Carga porcentual y valoración de los criterios de rigor científico en el IERCITEM.*

Criterio	Cantidad de ítems	%	Valoración/Puntuación				
			5	4	3	2	1
1. Metodológico	22	21.1	110	88	66	44	22
2. Axiológico	12	11.5	60	48	36	24	12
3. Praxiológico	7	6.8	35	28	21	14	7
4. Epistemológico	22	21.1	110	88	66	44	22
5. Ontológico	27	26.0	135	108	81	54	27
6. Aspectos formales	14	13.5	70	56	42	28	14
Total	104	100.0	520	416	312	208	104

Nota. Elaboración propia. 2025

Los seis criterios comprenden 33 indicadores, 104 ítems y 155 subítems o elementos. El sistema otorga el mayor peso al criterio ontológico (26.0 %), seguido por los criterios metodológico y epistemológico, con 21.1% cada uno, lo que refleja adecuadamente su centralidad en la investigación.

Las columnas sombreadas (4 y 5) muestran la zona de aprobación, con puntuaciones mínima y máxima en la que cada criterio se considera aprobado. De la misma manera, da cuenta de las puntuaciones totales de aprobación del informe de tesis dentro de la zona 4-5, desde 416/520 puntos (80.0%) hasta 520/520 puntos (100.0%). Para establecer el umbral mínimo de aprobación de una tesis, se partió del hecho de que las principales instituciones de enseñanza superior de la República Dominicana establecen el corte mínimo de 80.0/100.0 puntos para aprobar la defensa de una tesis de maestría.

Proceso de validación del IERCITEM

En la validación del IERCITEM se siguió el esquema descrito en *Ruta metodológica*. Para evaluar la validez de contenido, se tomaron en cuenta tres elementos: la pertinencia, la claridad de la redacción y la suficiencia de los ítems (Díaz Herrera, 2022). En la pertinencia se valoró la relevancia y la alineación de los ítems con los componentes correspondientes al constructo y a su redacción (Martínez-Valdés y Juárez-Hernández, 2019). Para la claridad se valoró si el ítem era de fácil comprensión y si su sintaxis y semántica eran adecuadas. En la suficiencia se valoró si la cantidad de ítems que pertenecían a un mismo criterio eran suficientes, según la escala 1-5.

La fiabilidad, por su parte, se centró en la coherencia con los análisis de α de Cronbach, el índice *CITC* (Corrected Item–Total Correlation), el índice de homogeneidad como también, en el análisis factorial.

Tanto la validez como la fiabilidad se determinaron desde un enfoque multimétodo, integrando técnicas estadísticas y lógicas clásicas con más avanzadas propias de cada criterio. El resultado fue un análisis robusto desde varias perspectivas que respaldó la obtención de conclusiones sólidas.

La selección de los expertos se hizo siguiendo la calificación del Coeficiente de Competencia (K) propuesto por Mendoza Fernández (2015). Los expertos seleccionados obtuvieron una media de Coeficiente de Competencia (K) = 0.87 (rango $0.8 < K < 1.0$) con una calificación de “Alto” (Marín-González et al., 2021).

Para el análisis, a los expertos se les entregó el instrumento de evaluación acompañado de un instructivo procedimental, una síntesis de los componentes esenciales de la investigación, una matriz de valoración tipo Likert (Joshil et al., 2015), y el protocolo de valoración de cada ítem y de cada criterio.

El análisis de normalidad evidenció que el 83.0% de los datos presentaban una distribución no paramétrica, al exceder los criterios de sesgo y curtosis establecidos por West et al., 1995. Por esta razón, los análisis estadísticos se realizaron bajo el supuesto de no normalidad.

No obstante, se utilizaron coeficientes de fiabilidad y concordancia debido a su reconocida robustez ante desviaciones moderadas de la normalidad en muestras de tamaño moderado y escalas compuestas. Asimismo, los resultados fueron corroborados mediante estadísticos no paramétricos complementarios, como la *rho* de Spearman y la *W* de Kendall, lo que fortaleció la consistencia y validez de las conclusiones obtenidas.

En el análisis estadístico de los datos, se utilizaron los programas EduG y G-String IV, el paquete IBM SPSS Statistics 24, Excel en sinergia con Python y la hoja de cálculo Excel. En la automatización de cálculos repetitivos y en la generación de informes vinculados a Excel se utilizaron las librerías de Python (pandas, openpyxl y xlswriter).

Para las visualizaciones gráficas se utilizaron la hoja de cálculo Excel 365/2021 y la asistencia de IA ChatGPT: modelos o3 y o5, (OpenAI, 2025) y Microsoft Copilot (IA).

Resultados

Validez del instrumento

Para garantizar la validez del IERCITEM se empleó una triangulación metodológica con índices reconocidos en la literatura especializada que abordan las tres dimensiones clave del juicio experto: pertinencia, claridad y suficiencia.

La Tabla 2 muestra los resultados de los ocho índices estadísticos que se utilizaron para evaluar la validez de contenido, estos fueron: el CVR de Lawshe (1975), el CVR' de Tristán (2008), el I-CVI, el S-CVI/AVE, el S-CVI/UA, la V de Aiken, su banda de confianza y el Índice de Congruencia de Expertos (ICE).

Tabla 2

Estadística de la validez de contenido del IERCITEM a partir de su pertinencia, claridad y suficiencia.

Criterios	CVR (Lawshe, 1975)	CVR' (Tristán, 2008)	I- CVI	S- CVI/ AVE	S- CVI/ UA	V Global (Aiken)	Banda de Aiken estimada	ICE
Pertinencia	0.95	0.97	0.98	0.97	0.93	0.94	0.93 - 0.96	0.95
Claridad	0.93	0.95	0.97	0.96	0.91	0.92	0.91 - 0.94	0.93
Suficiencia	0.91	0.93	0.95	0.94	0.89	0.91	0.90 - 0.92	0.92

Nota. Análisis realizado a partir de las evaluaciones inter-expertos al IERCITEM. Elaboración propia. 2025.

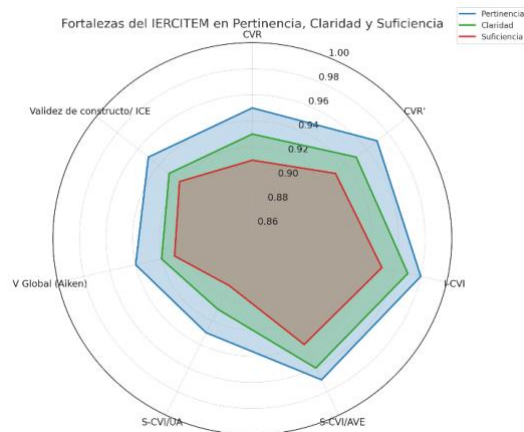
El IERCITEM exhibe una alta pertinencia en sus ítems, evidenciada por índices sobresalientes del *CVR* (0.95), *CVR'* (0.97), *I-CVI* (0.98), *S-CVI/AVE* (0.97) y *S-CVI/UA* (0.93), lo que indica un sólido consenso entre los jueces sobre su alineación conceptual y contextual. La *V* Global de Aiken (0.94) y su banda estimada (0.93–0.96), junto con el *ICE* (0.95), respaldan la validez estadística y metodológica del contenido.

En cuanto a la claridad y suficiencia, el IERCITEM también presenta indicadores sólidos. La claridad semántica y sintáctica alcanza valores elevados en $CVR = 0.93$, $I-CVI = 0.97$, V de Aiken = 0.92, y con $ICE = 0.93$, lo que garantiza la comprensión precisa del constructo evaluado. Aunque el $S-CVI/UA = 0.91$ muestra una ligera disminución, se mantiene dentro del rango óptimo. Por su parte, la suficiencia refleja una cobertura representativa del contenido ($CVR = 0.91$; $S-CVI/AVE = 0.94$; $ICE = 0.92$), con una banda de Aiken entre 0.90 y 0.92.

La Figura 1 muestra de forma visual la consistencia metodológica del instrumento en cada dimensión y evidencia valores que se mantienen por encima de los estándares internacionales de aceptabilidad.

Figura 1

Fortalezas de validez del IERCITEM en las dimensiones Pertinencia, Claridad y Suficiencia



Nota. Elaboración propia con asistencia de inteligencia artificial (ChatGPT, 2025).

Fiabilidad del instrumento

Para evaluar la fiabilidad del instrumento, se analizaron la consistencia interna (α de Cronbach), el grado de homogeneidad inter-ítems, la estabilidad temporal (ICC y correlación de Spearman, ρ), y la concordancia interjueces (W de Kendall y V de Aiken). Estos estadísticos se reforzaron con los índices SEM_{α} , la Teoría de la Generalizabilidad (Coeficiente G) y, de manera adicional, con la varianza de Índice de Facilidad (VIF).

La Tabla 3 detalla los resultados del cálculo de la fiabilidad a partir de los índices psicométricos asumidos. El análisis se centra en la fiabilidad interna, la estabilidad estadística y la concordancia interevaluadores, identificando las fortalezas, la consistencia entre ítems, y la solidez de cada dimensión evaluada en el instrumento. Además, se destaca el nivel de discriminación, la tolerancia entre ítems y la generalizabilidad del constructo, elementos clave para validar científicamente la estructura del instrumento.

Tabla 3

Evaluación de fiabilidad del IERCITEM.

Aspectos	Análisis estadístico	Resultado Global/ Promedio	Valoración
Consistencia interna	Alfa de Cronbach (α)	0.91	Muy alta
	Error estándar del Alfa (SEM_{α})	0.18	Bajo

	CITC promedio (Corrected Item- Total Correlation)	0.70	Buena discriminación
	Coeficiente G	0.85	Alta
Estabilidad temporal	Spearman (ρ)	0.82	Asociación fuerte
	ICC (General)	0.87	Muy alta
	ICC (2,1)	0.85	Muy alta fiabilidad
	ICC (2 K)	0.88	Buena
Concordancia interevaluadores	Spearman (ρ)	0.82	Asociación fuerte
	Kendall W	0.80	Alta concordancia
	ICC (General)	0.87	Muy alta
	ICC (2,1)	0.85	Muy alta fiabilidad
	ICC (2 K)	0.88	Buena
	Coeficiente G	0.85	Alta

Nota. Datos provenientes de la aplicación piloto del IERCITEM. 2025.

El IERCITEM muestra una fiabilidad psicométrica sobresaliente en las tres dimensiones clave: consistencia interna, estabilidad temporal y concordancia interjueces. El alfa de Cronbach ($\alpha = 0.91$) indica una fiabilidad interna muy alta, respaldada por una media de correlaciones ítem-total corregidas ($CITC = 0.70$), lo que evidencia una buena discriminación entre los ítems (Frías-Navarro, 2022). En cuanto a la estabilidad temporal, el coeficiente de correlación de Spearman ($\rho = 0.82$) y el ICC (2, $K = 0.88$) reflejan una asociación fuerte y una fiabilidad cercana a la excelencia (Koo y Li, 2016).

El análisis factorial exploratorio (AFE) con extracción de componentes principales y rotación varimax mostró una alta consistencia interna en todos los criterios, con coeficientes de alfa de Cronbach superiores a 0.87 y valores de Omega de McDonald ligeramente superiores en los modelos congénicos. Los criterios Metodológico, Epistemológico y Ontológico, presentaron cargas factoriales altamente homogéneas (≥ 0.82), mientras que los criterios Axiológico, Praxiológico y Formales presentaron cargas más heterogéneas (rango: 0.68–0.89).

Para la validez criterial, se tomó como referente el instrumento *Parámetro para Evaluar el Rigor Científico de las Tesis de Maestría en Educación* de Chambi Mescco (2017), que se aplicó a 72 tesis, lo que aportó datos de campo. Consta de 44 indicadores distribuidos en cinco dimensiones, y fue diseñado específicamente para evaluar el rigor científico de las tesis de maestría en Educación. Su validación fue la de Validez de contenido por juicio de expertos con una V de Aiken > 0.80 en todos los indicadores y se reporta consistencia inter-jueces validada.

El criterio de validación fue el concurrente y se asumió, no como *Gold Stándar*, sino como “criterio comparativo externo”. Para los criterios exclusivos del IERCITEM, se realizaron análisis complementarios que se detallan en la Tabla 3. Se utilizaron los datos procedentes de la aplicación piloto y se analizaron los siguientes aspectos de congruencia: correlación total de puntuaciones, correlación por dimensiones, concordancia en niveles de evaluación, concordancia continua por tesis, coincidencia en la clasificación cualitativa, y la dispersión de niveles por dimensiones. En la Tabla 4 se describen e interpretan los resultados de los aspectos analizados a la luz de los estadísticos aplicados en cada caso.

Tabla 4*Resultados comparativos de validez criterial entre IERCITEM y Chambi Mescco.*

Aspectos	Estadístico aplicado	Resultado	Interpretación
Correlación total de puntuaciones	Pearson (r)	$r = 0.86$	Alta correlación lineal
Correlación por dimensiones	Spearman (ρ)	$\rho = 0.80$ (Promedio)	Asociación ordinal fuerte entre las dimensiones homologadas
Concordancia en niveles de evaluación	Kappa de Fleiss	$\kappa = 0.72$	Buena coincidencia categórica
Concordancia continua por tesis	Coeficiente de Correlación Intraclass (ICC)	$ICC = 0.89$	Concordancia excelente entre puntuaciones totales
Coincidencia categórica	% de coincidencia categórica	78.4%	Alto. Ambos instrumentos presentan una fuerte correspondencia en la clasificación de los ítems (Polit y Beck, 2021)
Dispersión de niveles por dimensiones	W de Kendall	$W = 0.77$	Alta consistencia ordinal en las dimensiones evaluadas

Nota. Tratamiento estadístico de los resultados al aplicarles a las tesis los instrumentos IERCITEM y el de Chambi Mescco. 2025.

El IERCITEM exhibe una validez criterial sólida, ya que muestra una alta correspondencia con el referente evaluativo de Chambi Mescco. Desde el enfoque cuantitativo, se observan correlaciones significativas ($r \approx 0.86$; $\rho = 0.80$) entre las puntuaciones totales, así como una excelente concordancia en las dimensiones evaluadas ($ICC > 0.85$; $W = 0.77$), lo que indica que el comportamiento del instrumento es consistente y reproducible frente al criterio establecido.

Desde la perspectiva categórica, se observa una coincidencia del 78.4% en los niveles evaluativos, respaldada por un coeficiente kappa de Fleiss de 0.72, lo que confirma la congruencia cualitativa entre ambos instrumentos. Estos resultados constituyen una evidencia empírica sólida que respalda el uso del IERCITEM como herramienta legítima para evaluar el rigor científico de las tesis de maestría en educación, tanto en sus dimensiones operativas como interpretativas.

Discusión y conclusiones

La convergencia de los distintos índices empleados en el proceso de validación aporta evidencia consistente de que la estrategia metodológica adoptada, sustentada en

el juicio de expertos, el análisis de contenido y la validación empírica, fue adecuada para respaldar la validez y fiabilidad del instrumento. Particularmente, la validación por juicio de expertos evidenció una elevada correspondencia entre el constructo teórico y los contenidos representados en los ítems, lo que fortalece la consistencia conceptual del IERCITEM.

Estos resultados adquieren especial relevancia cuando se interpretan a la luz de las deficiencias identificadas en investigaciones sobre tesis de maestría, las cuales han documentado debilidades metodológicas, inconsistencias teóricas y limitaciones en la originalidad científica (Francia Hernández, 2023; Ge et al., 2024; Huayta Meza, 2021; Nina Condori, 2022). En este contexto, el IERCITEM constituye una propuesta sustentada empíricamente para contribuir a la evaluación sistémica del rigor científico en investigaciones de posgrado.

La comparación con propuestas previas permite identificar aportes diferenciales relevantes. Los trabajos de Aliaga-Pacora et al. (2023), Ceballos-Almeraya y Tobón (2019) y Cabrera et al. (2023) fueron concebidos principalmente para valorar productos académicos específicos. Aunque dichas propuestas han demostrado utilidad en sus respectivos ámbitos de aplicación, el IERCITEM se orienta hacia la evaluación del rigor científico como un constructo multidimensional, integrando las dimensiones metodológicas, epistemológicas, ontológicas, axiológicas, praxiológicas y aspectos formales en un mismo marco de análisis.

Se supera, de esta manera, el alcance evaluativo tradicional y le confiere al instrumento una utilidad adicional como referente orientador para el desarrollo de investigaciones desde sus etapas iniciales en contextos locales e internacionales.

Los hallazgos respaldan la propuesta teórica de Deroncele Acosta et al. (2021), quienes conciben el rigor científico como un fenómeno multidimensional constituido por componentes metodológicos, epistemológicos, ontológicos, axiológicos y praxiológicos interrelacionados. La elevada consistencia observada entre los criterios del IERCITEM aporta evidencia empírica favorable a esta conceptualización y demuestra la posibilidad de operacionalizar tales dimensiones con un instrumento sometido a procesos sistemáticos de validación psicométrica.

Esta condición resulta coherente con los planteamientos de Conner et al. (2025), quienes sostienen que los constructos complejos deben representarse a través de dimensiones conceptualmente diferenciadas, pero empíricamente vinculadas.

Asimismo, las evidencias obtenidas de validez y fiabilidad sitúan al IERCITEM en consonancia con los estándares contemporáneos de validación psicométrica, fortaleciendo su pertinencia para contextos de evaluación académica. En este sentido, los resultados sugieren que el instrumento posee condiciones adecuadas para ser utilizado en procesos de valoración del rigor científico en tesis de maestría en educación.

Desde otra perspectiva, Guillén Díaz y Sanz Trigueros (2021) argumentan que la calidad de la investigación trasciende el rigor metodológico e incorpora elementos relacionados con la relevancia, la contribución al conocimiento y el impacto científico. Aunque el IERCITEM fue diseñado para evaluar el rigor científico, su estructura incorpora dimensiones éticas, ontológicas y formales que habitualmente reciben menor atención en los procesos convencionales de evaluación de tesis. Por ello, puede contribuir a una valoración más amplia de aspectos asociados a la calidad académica de la investigación, sin perder su foco principal en la evaluación del rigor científico.

Los resultados también permiten situar al IERCITEM dentro del conjunto de instrumentos destinados a evaluar constructos educativos complejos (Castillo-de-León et al., 2024; Conner et al., 2025; Hartelt y Martens, 2025). Sin embargo, su principal

particularidad radica en la integración de una finalidad evaluativa y orientadora centrada específicamente en el rigor científico de las investigaciones educativas. Esta característica amplía las posibilidades de aplicación del instrumento en procesos de aseguramiento de la calidad del posgrado.

Finalmente, la comparación con la propuesta de Chambi Mescco (2017) permite apreciar tanto convergencias como diferencias sustantivas. Ambos instrumentos comparten el propósito de valorar el rigor científico en tesis de maestría y presentan fundamentos conceptuales compatibles. No obstante, mientras la propuesta de Chambi Mescco se concentra principalmente en la evaluación del informe de investigación, el IERCITEM incorpora criterios distintivos que amplían la comprensión del proceso de construcción del conocimiento.

La diferencia entre el IERCITEM y el instrumento usado como criterio fortalece su capacidad para valorar de manera más comprehensiva las diversas dimensiones que intervienen en la producción científica dentro del ámbito educativo.

Conclusiones

El IERCITEM, con una validez y fiabilidad sobresalientes, se presenta como una herramienta rigurosa para evaluar la calidad científica de las tesis de maestría en educación. Los análisis de validez de contenido arrojan índices superiores a los estándares aceptados, lo que confirma la pertinencia, claridad y suficiencia de los ítems. La alta correlación en la validez criterial con el instrumento de referencia respalda empíricamente su eficacia orientativa a la investigación y evaluativa del producto. La fiabilidad global evidencia una notable consistencia interna y estabilidad temporal, reforzada por la alta concordancia interevaluadores.

Sin embargo, la validez exige evidencias acumulativas y contextuales, por lo que ningún resultado aislado resulta concluyente (Polit y Beck, 2021).

Los resultados del AFE confirman la estructura multidimensional y la solidez conceptual de sus seis criterios, distinguiendo adecuadamente entre dimensiones homogéneas y heterogéneas. En conjunto, estos resultados confirman que el IERCITEM es un instrumento válido, fiable y aplicable en contextos académicos similares al de la UASD-Recinto Barahona de la República Dominicana.

El objetivo de esta investigación, que consistió en diseñar un marco para la orientación y la evaluación del rigor científico en las tesis, se alcanzó en su plenitud. De este modo, se contribuye al aseguramiento de la calidad de las tesis de maestría en educación con altos estándares de rigor científico, lo que aumenta su credibilidad y garantiza la reproducibilidad de sus resultados.

Limitaciones del IERCITEM

La relativa homogeneidad geográfica e institucional del panel de expertos utilizado para la validación de contenido, pudo haber restringido la amplitud de perspectivas incorporadas al proceso de validación. Aunque los jueces poseían experiencia acreditada en investigación educativa, la mayoría procedía de contextos académicos similares, por lo que investigaciones futuras podrían incluir expertos de diferentes países e instituciones para fortalecer la generalización y robustez de la evidencia de validez.

El tamaño de la muestra piloto, adecuada para análisis iniciales, se considera insuficiente para estudios psicométricos de mayor complejidad o modelamientos multi grupales, como el análisis factorial confirmatorio, las estimaciones robustas de validez o la confiabilidad avanzada. Además, el predominio de datos no paramétricos limitó el uso de técnicas más robustas.

Finalmente, la validez criterial con un solo referente, la ausencia de triangulación cualitativa y de validación multi transcultural, junto con la necesidad de revisar los ítems de aquellos criterios con cargas heterogéneas, indican áreas de mejora para fortalecer su estructura, aplicabilidad y validez internacional.

Recomendaciones

Se recomienda fortalecer la validez de constructo del IERCITEM mediante la realización de un Análisis Factorial Confirmatorio (AFC) con una muestra más amplia y diversa, incorporando índices de ajuste avanzados (CFI, RMSEA, TLI, SRMR) y, eventualmente, modelos de ecuaciones estructurales.

Asimismo, se sugiere ampliar la aplicación interinstitucional a otras universidades dominicanas y latinoamericanas, con el fin de consolidar la validez externa e identificar patrones comparativos del rigor científico en diferentes contextos educativos.

Igualmente, se plantea el desarrollo de una versión digital, o en línea, del instrumento para facilitar la aplicación automatizada, la recogida de datos y la generación de informes. Además, se propone su validación transcultural en inglés y portugués para fomentar estudios comparativos internacionales

Por último, se recomienda implementar programas de capacitación para asesores y evaluadores de tesis, incorporar análisis cualitativos como entrevistas o grupos de discusión, y realizar una revisión periódica del instrumento.

Agradecimientos

Se agradece la colaboración de los asesores, jurados examinadores de tesis de maestría en educación, expertos y evaluadores del instrumento, cuya valiosa contribución en la recopilación de información y en la validación del instrumento ha sido fundamental para su mejora. Este trabajo refleja el compromiso compartido con el fortalecimiento de la investigación educativa en la República Dominicana.

Conflicto de intereses

Los autores declaran que no existe ningún conflicto de interés en torno a esta investigación. La misma se llevó a cabo con imparcialidad, transparencia y apego a la ética académica, y sus hallazgos se sustentan exclusivamente en la evidencia recopilada y en la evaluación objetiva de los datos.

Referencias

- Aliaga-Pacora, A. A., Juárez-Hernández, L. G., y Sumarriva-Bustinza, L. A. (2023). Validez de una rúbrica para evaluar competencias investigativas en postgrado. *Revista de Ciencias Sociales*, 29(4), 415-427. <https://doi.org/10.31876/rsc.v29i4.41265>
- Arguimbau, L. (2023). *La UOC, hacia una nueva evaluación científica*. Universitat Oberta de Catalunya. <https://www.uoc.edu/es/news/2023/237-la-UOC-impulsa-la-nueva-evaluacion-de-la-investigacion>
- Cabrera, P. R., Ruiz Bolívar, C., Gomes, T. P., y Beretta, R. L. (2023). Desarrollo de una escala para medir competencias investigativas en docentes y estudiantes universitarios. *Areté*, 9(17), 147-169. <https://doi.org/10.55560/arete.2023.17.9.7>

- Castillo-de-León, M. A., Méndez-Hinojosa, L. M., y Cárdenas-Rodríguez, M. (2024). Development and psychometric validation of an information competency assessment: The information management brief scale. *European Journal of Educational Research*, 13(2), 619-630. <https://doi.org/10.12973/eu-jer.13.2.619>
- Ceballos-Almeraya, J. M., y Tobón, S. (2019). Validez de una rúbrica para medir competencias investigativas en pedagogía desde la socioformación. *Atenas*, 3(August), 1-17. <https://www.redalyc.org/journal/4780/478060102001/html/>
- Chambi Mescco, E. (2017). Nivel de rigor científico de las tesis de maestría en Educación de la UNMSM desde el año 2012 al 2014. *Consensus*, 22(1), 37-47. <https://doi.org/10.33539/consensus.v22n1.991>
- Conner, J., Holquist, S. E., Mitra, D. L., y Boat, A. (2025). Measuring student voice practices: Development and validation of school and classroom scales. *AERA Open*, 11, 1-24. <https://doi.org/10.1177/23328584251326893>
- Corral de Franco, Y. (2022). Validez y confiabilidad en instrumentos de investigación: una mirada teórica. *Revista Ciencias de la Educación*, 32(60), 562-586. <https://servicio.bc.uc.edu.ve/educacion/revista/60/art06.pdf>
- Correa Hincapié, N., Cañas Gutiérrez, A. I., Cogollo Flórez, J. M., Romero Sáez, M., Jaramillo Osorio, A. F., y Zuluaga, R. (2023). Calidad científica: Definición, historia y aplicaciones. *SIGNOS – Investigación en Sistemas de Gestión*, 15(1). <https://doi.org/10.15332/24631140.8247>
- Deroncele Acosta, A., Gross Tur, R., y Medina Zuta, P. (2021). El mapeo epistémico: herramienta esencial en la práctica investigativa. *Revista Universidad y Sociedad*, 13(3), 172-188. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2218-36202021000300172&lng=es&tlng=es
- Díaz Herrera, L. C. (2022). Validez (Capítulo 3). En Pedro A. Sánchez Escobedo (Ed.), *Diseño y desarrollo de Instrumentos en Línea* (pp. 37- 47). Centro de Estudios, Clínica e Investigación Psicoanalítica S.C. https://www.researchgate.net/publication/365392882_Diseño_y_desarrollo_de_Instrumentos_en_Línea
- Díaz Torres, J. H. (2024). La investigación y las tesis. A diez años de la Ley Universitaria en el Perú (2014-2024). *Yachay - Revista Científico Cultural*, 13(2), 85-95. <https://doi.org/10.36881/yachay.v13i2.945>
- Dunn, T. J., Baguley, T., y Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399-412. <https://doi.org/10.1111/bjop.12046>
- Escurra Mayaute, L. (1988). Cuantificación de la validez de contenido por criterio de jueces. *Revista de Psicología*, 6(1-2), 103-111. <https://doi.org/10.18800/psico.198801-02.008>
- Francia Hernández, M. E. (2023). *Rigor científico de los proyectos de investigación de maestría en la escuela de posgrado, facultad de Ciencias y Humanidades, Universidad de El Salvador construcción* [Tesis de Maestría, Universidad de El Salvador]. <https://repositorio.ues.edu.sv/server/api/core/bitstreams/0c50e035-7e7f-48bf-ad02-6aec4beb1c66/content>
- Francia, M. E., y Figueroa Morán, H. E. (2022). Importancia del aparato crítico en proyectos de investigación de maestrías en la Universidad de El Salvador. *Revista Diálogo Interdisciplinario sobre Educación*, 4(2), 101-112. <https://revistas.ues.edu.sv/index.php/redised/article/view/2781/2766>

- Ge, Z.-G., Zhao, Y., y Liu, Y. (2024). Exploring quality issues in academic master theses of education majors and corresponding countermeasures. *Journal of Pedagogical Research*, 8 (4), 30-47. <https://doi.org/10.33902/JPR.202428171>
- González Martínez, M. de la C., Keeling Alvarez, M., y Pérez Lemus, L. (2022). Contribuição científica na pesquisa educacional: algumas considerações. *Revista Angolana De Ciencias*, 4(2), e040202. <https://doi.org/10.54580/R0402.02>
- Guillén Díaz, C., y Sanz Trigueros, F. J. (2021). El rigor científico en investigación. Consideraciones desde el área de Didáctica de la Lengua y la Literatura. *El Guiniguada*, 30, 40–51. <https://doi.org/10.20420/ElGuiniguada.2021.402>
- Hartelt, T., y Martens, H. (2025). Development and validation of an instrument measuring students' self-efficacy regarding explaining evolutionary changes. *Journal of Biological Education*, 1–35. <https://doi.org/10.1080/00219266.2025.2546795>
- Huayta Meza, M. E. (2021). *Calidad metodológica de los trabajos de grado de la maestría en educación de una escuela de posgrado, Perú. Período 2010 – 2019*. Universidad Peruana Unión [Tesis de maestría en Educación con mención en investigación y docencia universitaria]. <https://repositorio.upeu.edu.pe/server/api/core/bitstreams/ee9cc029-d7af-47bc-b5c5-fc514613731f/content>
- Joshil, A., Kale, S., Chandel, S. y Pal, D. K. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science and Technology*, 7(4), 396-403. <https://doi.org/10.9734/BJAST/2015/14975>
- Koo, T. K., y Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Lawshe, Ch. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Martínez-Arias, M. R. (2005). *Psicometría: Teoría de los tests psicológicos y educativos*. Síntesis. https://indaga.ual.es/discovery/fulldisplay/alma991000243049704991/34CBUA_UAL:VU1
- Martínez-Valdés, M. G., y Juárez-Hernández, L. G. (2019). Diseño y validación de un instrumento para evaluar la formación en sostenibilidad en estudiantes de educación superior. *IE Revista de Investigación Educativa de La REDIECH*, 10(19), 37 https://doi.org/10.33010/ie_rie_rediech.v10i19.501
- McDonald, R. P. (1999). *Test Theory: A Unified Treatment* (1st ed.). Psychology Press. <https://doi.org/10.4324/9781410601087>
- Mendoza Fernández, S. H. (2015). *Criterio de expertos. Su procesamiento a través del método Delphi*. Universidad de Barcelona. <https://www.histodidactica.com/criterio-de-expertos-su-procesamiento-a-traves-del-metodo-delphi/>
- Nina Condori, M. (2022). *Análisis de las tesis de maestría en educación superior UNSAAC, 2008-2020*. Universidad Nacional San Antonio Abad del Cusco [Tesis de maestría en Educación Superior, Universidad Nacional San Antonio Abad del Cusco]. https://repositorio.unsaac.edu.pe/bitstream/handle/20.500.12918/7358/253T20_221196_TC.pdf?sequence=1&isAllowed=y
- Ochoa Sierra, L., y Moreno Mosquera, E. (2019). Estructura y movidas de la sección “metodología” en tesis de posgrado de educación. *Enunciación*, 24(2), 133-151. <http://doi.org/10.14483/22486798.14772>
- Polit, D. F., y Beck, C. T. (2021). *Nursing Research: Generating and Assessing Evidence for Nursing Practice* (11th ed.). Wolters Kluwer.

- Soto, C. (2021). Development and validation of an instrument to measure the methodological quality of master's and doctoral theses. *Revista Internacional de Investigación en Ciencias Sociales*, 17(2), 357-378. [10.18004/riics.2021.diciembre.357](https://doi.org/10.18004/riics.2021.diciembre.357)
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72-101. <https://doi.org/10.2307/1412159>
- Tristán-López, A. (2008). Modificación al modelo de Lawshe para el dictamen cuantitativo de la validez de contenido de un instrumento objetivo. *Avances en Medición Instituto de Evaluación e Ingeniería Avanzada*, 6(1), 37-48. <https://dialnet.unirioja.es/servlet/articulo?codigo=2981185>
- Vásquez-Antonio, J. M., Tobón, S., Vásquez-Antonio, J., Herrera-Meza, S. R., y Juárez-Hernández, L. G. (2018). Diseño y validez de contenido de una rúbrica socioformativa para evaluar el informe de prácticas profesionales en la Educación Normal. *Revista Espacios*, 39(53) <https://www.revistaespacios.com/cited2017/cited2017-29.pdf>
- West, S. G., Finch, J. F., y Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56-75). Sage Publications, Inc. <https://psycnet.apa.org/record/1995-97753-004>

Anexo.

Prompt:

Figura 1

Diseña una figura científica titulada “Fortalezas de validez del IERCITEM en las dimensiones Pertinencia, Claridad y Suficiencia”. Utiliza un gráfico radial con zonas diferenciadas por colores tenues, en estilo académico para publicación científica. Presenta tres dimensiones: Pertinencia, Claridad y Suficiencia. Incluye los siguientes valores:

- Pertinencia: $CVR=0.95$, $CVR'=0.97$, $I-CVI=0.98$, $S-CVI/AVE=0.97$, $S-CVI/UA=0.93$, V de Aiken= 0.94 , $ICE=0.95$.
- Claridad: $CVR=0.93$, $CVR'=0.95$, $I-CVI=0.97$, $S-CVI/AVE=0.96$, $S-CVI/UA=0.91$, V de Aiken= 0.92 , $ICE=0.93$.
- Suficiencia: $CVR=0.91$, $CVR'=0.93$, $I-CVI=0.95$, $S-CVI/AVE=0.94$, $S-CVI/UA=0.89$, V de Aiken= 0.91 , $ICE=0.92$.

