



## VALIDEZ DE CONTENIDO MEDIANTE JUICIO DE EXPERTOS EN RÚBRICAS EN EL CONTEXTO DE UN LABORATORIO SOBRE TEMAS DE MAGNETISMO

## VALIDITY OF CONTENT BY EXPERT JUDGEMENT IN RUBRICS IN THE CONTEXT OF A LABORATORY ON MAGNETISM TOPICS

## VALIDADE DO CONTEÚDO POR AVALIAÇÃO DE ESPECIALISTAS EM RUBRICAS NO CONTEXTO DE UM LABORATÓRIO SOBRE TEMAS DE MAGNETISMO

Vladimir Camelo Avedoy\*<sup>ID</sup>, Mario Humberto Ramírez Díaz\*\*<sup>ID</sup>  
José Luis Santana Fajardo\*\*\*<sup>ID</sup>, Abigail Rojas-Glez\*\*\*\*<sup>ID</sup>

Camelo, V., Ramírez, M., Santana, J., Rojas-Glez, A. (2026). Validez de contenido mediante juicio de expertos en rúbrica en el contexto de un laboratorio sobre temas de magnetismo. *Góndola, Enseñanza y Aprendizaje de las Ciencias*, 21(1), pp. 1-11 e23292. <https://doi.org/10.14483/23464712.23292>

### Resumen

En el campo de la física educativa, el uso de la taxonomía funcional de Ribes para el diseño de rúbricas es poco frecuente, especialmente en el área de magnetismo, donde no se reportan antecedentes de validación de tales instrumentos. El presente trabajo tiene como objetivo validar una rúbrica diseñada bajo los cinco niveles funcionales de dicha taxonomía: contextual, suplementario, selector, sustitutivo referencial y sustitutivo no referencial. El instrumento se desarrolló para evaluar el nivel de competencias en estudiantes de ingeniería que cursan el laboratorio de electromagnetismo en el Centro Universitario de Ciencias Exactas e Ingenierías de la Universidad de Guadalajara, México. La validez de contenido se determinó mediante juicio de expertos, evaluando pertinencia, claridad conceptual, niveles cognoscitivos, formato, redacción y terminología a través de una escala Likert de cinco puntos. Metodológicamente, se aplicó el procedimiento de Hernández-Nieto para el cálculo del Coeficiente de Validez de Contenido (CVC), siguiendo el protocolo implementado en investigaciones recientes sobre ciclos de aprendizaje. Los resultados obtenidos arrojaron valores de CVC situados entre 0.90963 y 0.95496. Estos indicadores se categorizan como "excelentes" para todos los niveles de complejidad de la taxonomía propuesta. Se concluye que el instrumento es válido y

\* Maestro en Ciencias Física. Profesor titular A en Departamento de Física del CUCEI, U. de G., México. ORCID: <https://orcid.org/0009-0009-0119-3741>

\*\* Doctor en Ciencias en Física Educativa. Profesor titular C en CICATA-LEGARIA-IPN, México. ORCID: <https://orcid.org/0000-0002-3459-2927>

\*\*\* Doctor en Ciencias en Física Educativa. Profesor asociado C en departamento de Física del CUCEI, U. de G., México. ORCID: <https://orcid.org/0000-0001-7048-2648>

\*\*\*\* Master of Philosophy, Profesor titular C en Departamento de Física del CUCEI, U. de G., México. [abigail.rojas@academicos.udg.mx](mailto:abigail.rojas@academicos.udg.mx) – ORCID: <https://orcid.org/0000-0003-2051-8328>

pertinente para la evaluación de competencias en el contexto experimental del magnetismo.

**Palabras Clave:** Rúbrica; competencias; magnetismo; juicio de expertos; coeficiente de validez de contenido.

### **Abstract**

In the field of physics education, the application of Ribes' functional taxonomy for rubric design remains uncommon. Specifically, within the area of magnetism, there is a lack of published literature regarding the development and validation of such assessment instruments. This study aims to validate a rubric based on Ribes' five functional levels: contextual, supplementary, selector, referential substitutive, and non-referential substitutive. The instrument was designed to assess the level of competency development in undergraduate engineering students at the electromagnetism laboratory of the University Center of Exact Sciences and Engineering (CUCEI), University of Guadalajara, Mexico. Content validity was established through expert judgment, evaluating relevance, conceptual clarity, cognitive levels, format, wording, and terminology using a 5-point Likert scale. Methodologically, the Content Validity Coefficient (CVC) was calculated following Hernández-Nieto's procedure, a benchmark used in recent educational research. Results showed CVC values ranging from 0.90963 to 0.95496, which are classified as "excellent" across all complexity levels of the taxonomy. These findings demonstrate that the rubric is a highly valid and reliable tool for assessing competencies in magnetism laboratory settings.

**Keywords:** Rubric; skills; magnetism; expert judgment; content validity coefficient

### **Resumo**

No campo da física educativa, o uso da taxonomia funcional de Ribes para o desenho de rubricas é pouco frequente, especialmente na área do magnetismo, onde não se encontram antecedentes de validação de tais instrumentos. O presente trabalho tem como objetivo validar uma rubrica desenvolvida com base nos cinco níveis funcionais desta taxonomia: contextual, suplementar, seletor, substitutivo referencial e substitutivo não referencial. O instrumento foi concebido para avaliar o nível de competências de estudantes de engenharia no laboratório de eletromagnetismo do Centro Universitário de Ciências Exatas e Engenharia da Universidade de Guadalajara, México. A validade de conteúdo foi determinada mediante julgamento de especialistas, avaliando pertinência, clareza conceitual, níveis cognitivos, formato, redação e terminologia através de uma escala Likert de cinco pontos. Metodologicamente, aplicou-se o procedimento de Hernández-Nieto para o cálculo do Coeficiente de Validade de Conteúdo (CVC). Os resultados obtidos apresentaram valores de CVC entre 0,90963 e 0,95496, categorizados como "excelentes" para todos os níveis de complexidade da taxonomia proposta. Conclui-se que o instrumento possui alta validade e confiabilidade para a avaliação de competências no contexto experimental do magnetismo.

**Palavras-Chave:** Rubrica; competências; magnetismo; julgamento de especialistas; coeficiente de validade de conteúdo

## 1. Introducción

En el ámbito de la física educativa, la evaluación del aprendizaje es un proceso sistemático que permite recolectar y analizar datos para fundamentar juicios sobre el progreso de los estudiantes (Sierra, 2023; Fernández, 2017). Tradicionalmente, este proceso se ha apoyado en taxonomías como las de Bloom, Marzano y Kendall, o el modelo SOLO, las cuales cuentan con un amplio recorrido en temas de electricidad y magnetismo (Venier et al., 2019; Garzón, 2023). No obstante, la taxonomía funcional de Ribes (Ribes, 1985; Ribes-Iñesta, 2008), que plantea que el comportamiento es resultado de la interacción de un sujeto con su medio en distintos niveles de complejidad, ha sido escasamente explorada en la enseñanza de las ciencias físicas.

A pesar de que González (2007) realizó una adaptación de esta taxonomía en cursos de electromagnetismo, el área específica de magnetismo carece de antecedentes robustos. En la Universidad de Guadalajara (U. de G.), bajo el modelo de formación por competencias —influenciado por el proyecto Tuning en América Latina (Ramírez, 2018)—, es imperativo contar con instrumentos que midan no sólo el conocimiento disciplinar, sino también las dimensiones metodológicas y actitudinales. En este contexto, surge la necesidad de diseñar e implementar matrices de evaluación (rúbricas) que permitan identificar el nivel de logro de los estudiantes en los laboratorios experimentales.

Sin embargo, el diseño de una rúbrica no es suficiente; su utilidad depende de su calidad técnica, específicamente de su validez y confiabilidad. Como señalan Medina-Díaz y Verdejo-Carrión (2020), en la práctica docente suele asumirse la validez de los instrumentos de forma tácita, omitiendo la evidencia técnica necesaria. La validez de contenido, definida como el juicio lógico sobre la correspondencia entre los rasgos del aprendizaje y lo incluido en la prueba (García, 2002), es un requisito indispensable para que un instrumento sea considerado confiable.

Para subsanar esta carencia, el presente trabajo se centra en la validación mediante juicio de expertos de una rúbrica diseñada bajo los cinco niveles funcionales de Ribes: contextual, suplementario, selector, sustitutivo referencial y sustitutivo no referencial. Para ello, se emplea el Coeficiente de Validez de Contenido (CVC) propuesto por Hernández-Nieto (2002; 2024). A diferencia de otros coeficientes como Kappa o Pearson, el CVC es un método robusto que mide simultáneamente la validez de contenido y el nivel de acuerdo entre jueces, permitiendo refinar los ítems mediante criterios de pertinencia, claridad conceptual y redacción (Sánchez, 2021).

Bajo este marco, el presente estudio busca dar respuesta a la siguiente interrogante: ¿en qué medida la rúbrica diseñada a partir de la taxonomía de Ribes evalúa con precisión el constructo de interés en el laboratorio de magnetismo? El objetivo principal es, por tanto, determinar la validez de contenido de dicho instrumento, proporcionando una herramienta técnica y rigurosa para la evaluación de competencias en ingeniería.

Los saberes que integran la competencia del Laboratorio de Electromagnetismo se presentan en la *Tabla 1*, organizados en tres dimensiones fundamentales: cognitiva (*conocimientos*), procedimental (*habilidades*) y actitudinal (*actitudes y valores*). La dimensión del *conocimiento* abarca las leyes fundamentales del magnetismo, desde la Ley de Biot-Savart hasta la inducción de Faraday. Por su parte, la dimensión de habilidades se centra en la capacidad analítica y la comunicación científica, mientras que el eje de actitudes y valores prioriza el trabajo colaborativo y la seguridad en el laboratorio.

### Tabla 1

*Saberes de programa de Laboratorio de Electromagnetismo*

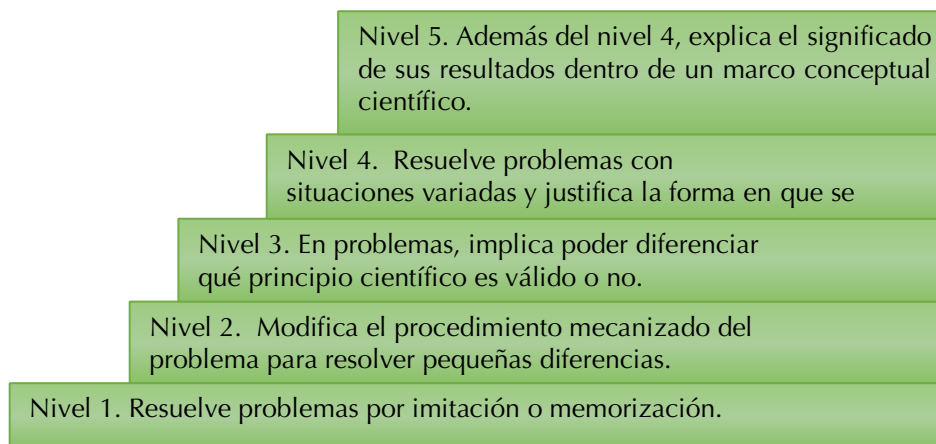
Conocimiento (Saber)	Habilidades (Saber hacer)	Actitudes y valores (Saber ser)
-Medición de campos magnéticos de bobinas y toroides.	-Analiza y relaciona los datos obtenidos.	-Trata con respeto a sus pares.
-Aplica interacción entre bobinas e imanes.	- Expresa, redacta y comunica con sustento científico.	-Propone y negocia para trabajar en equipo.
-Ley de Gauss del magnetismo y Ley de Biot- Savart.		-Hace buen uso del equipo y del mobiliario del laboratorio
-Ley de Ampere.		-.
-Ley de inducción de Faraday.		

Nota: Los saberes que integran la competencia del Laboratorio de Electromagnetismo se presentan organizados en tres dimensiones fundamentales: cognitiva (conocimientos), procedimental (habilidades) y actitudinal (actitudes y valores). Fuente: Programa del Laboratorio de Electromagnetismo aprobado en la Academia de Electromagnetismo del Departamento de Física del Centro universitario de Ciencias Exactas e Ingenierías, Universidad de Guadalajara.

De estos saberes se derivaron los indicadores que posteriormente fueron sometidos al juicio de expertos para validar su pertinencia.

### Figura 1

Niveles de complejidad alcanzados con base en la taxonomía de Ribes.



Nota: La operacionalización de los saberes descritos en la Tabla 1 se realizó mediante la escala jerárquica de la taxonomía de Ribes mostrada en la Figura 1 (Ribes e Iñesta, 1985). Fuente: Adaptado de Rodríguez, M. E. (2000, mayo). Hacia una taxonomía funcional del aprendizaje útil en la elaboración de exámenes objetivos [Ponencia]. Foro-taller de políticas y criterios para exámenes de academia, CUCEI, Universidad de Guadalajara, México.

Esta integración permitió que cada descriptor de la rúbrica no sólo evaluará la presencia del conocimiento, sino la complejidad funcional de la interacción del estudiante con el objeto de estudio.

A continuación, se describe la lógica de progresión utilizada para la construcción de los descriptores de desempeño:

1. *Nivel Contextual (1)*: El estudiante identifica o reproduce las leyes (ej. Ley de Biot-Savart) mediante la memoria o la imitación de un procedimiento estándar sin ajustes.
2. *Nivel Suplementario (2)*: El estudiante realiza mediciones y ajustes instrumentales básicos ante variaciones menores en el montaje del laboratorio, demostrando una adaptación procedimental.
3. *Nivel Selector (3)*: Ante un fenómeno magnético específico, el estudiante discrimina y elige la ley o el principio científico adecuado (ej. elegir entre la Ley de Ampere o la de Faraday) para abordar el problema.
4. *Nivel Sustitutivo Referencial (4)*: El estudiante resuelve problemas en contextos experimentales variados, justificando técnicamente sus decisiones y vinculando los datos obtenidos con la teoría física.
5. *Nivel Sustitutivo no Referencial (5)*: El estudiante trasciende la resolución del caso particular para explicar el comportamiento del sistema magnético dentro de un marco conceptual amplio, siendo capaz de proponer generalizaciones o reflexiones teóricas sobre los resultados.

Esta estructura asegura que la evaluación sea acumulativa; es decir, para que un estudiante alcance un desempeño de Nivel 4 (Sustitutivo Referencial) en el saber de "Medición de campos magnéticos", primero debe haber demostrado dominio en los niveles previos de identificación, ajuste y selección de principios.

Para ilustrar la transposición didáctica, la *Tabla 2* muestra el descriptor de desempeño diseñado para evaluar el dominio de la Ley de Ampere.

**Tabla 2**

*Ejemplo de indicadores de la rúbrica según niveles funcionales.*

Nivel Funcional	Descriptor de Desempeño (Indicador)
1. Contextual	Ejemplifica el concepto y las propiedades de la circulación de la intensidad del campo magnético en un contorno cerrado en proporción a la corriente que lo atraviesa.
2. Suplementario	Utiliza el concepto y propiedades de la circulación de la intensidad del campo magnético en un contorno cerrado en proporción a la corriente que lo atraviesa de un ejemplo que ha sido empleado en clase y se ha cambiado los valores o datos proporcionados, pero sin alterar la secuencia del procedimiento.
3. Selector	Soluciona una situación problema empleando criterios de validez contenidos en el concepto y propiedades de la circulación de la intensidad del campo magnético en un contorno cerrado en proporción a la corriente que lo atraviesa.
4. Sustitutivo Referencial	Resuelve una situación problema que ha sido modificada utilizando la relación de la corriente que atraviesa un contorno cerrado con la intensidad del campo magnético, así como sus conceptos y propiedades introduciendo nuevos elementos que son necesarios.
5. Sustitutivo no Referencial	Resuelve una situación problema real o simulada juzgando y justificando el uso de concepto y propiedades de la circulación de la intensidad del campo magnético en un contorno cerrado en proporción a la corriente que lo atraviesa. Introduce nuevos elementos de otras disciplinas científicas necesarias para resolverlo.

Nota: Descriptor de desempeño diseñado para evaluar el dominio de la Ley de Ampere. Fuente: Elaboración propia.

La redacción de estos descriptores permitió que los jueces expertos evaluaran no solo si el contenido técnico era correcto, sino si la progresión de la complejidad funcional era coherente con los objetivos de aprendizaje del laboratorio. Esta estructura jerárquica facilita al docente identificar el 'techo' de desempeño de cada estudiante, permitiendo una evaluación más precisa que la asignación de un puntaje tradicional.

Para ver rúbrica completa consultar:

[https://drive.google.com/file/d/1U14LvWR4rsYIMPTaywhe-psA7SL7H\\_Zy/view?usp=sharing](https://drive.google.com/file/d/1U14LvWR4rsYIMPTaywhe-psA7SL7H_Zy/view?usp=sharing)

## 2. Metodología

Para validar la matriz de evaluación y determinar la pertinencia y relevancia de sus ítems, se empleó el método propuesto por Hernández-Nieto 2002; 2024, el cual consiste en la revisión por parte de los expertos para que un instrumento apruebe ciertos requisitos y se tenga mayor seguridad de que el instrumento aplicado sea el adecuado (Sánchez, 2021).

El análisis de los datos se realizó mediante software de hoja de cálculo, aplicando las ecuaciones de Hernández-Nieto (2024). El *CVC* se determinó promediando las valoraciones de los jueces por ítem y aplicando la corrección por concordancia aleatoria mediante las siguientes expresiones:

$$CVC_i = \frac{M_x}{V_{max}} \quad (1)$$

donde  $M_x$  corresponde al valor promedio de los jueces expertos por nivel y  $V_{max}$  el valor máximo que puede alcanzar la escala. Adicional a ello, es necesario calcular el error asignado a cada nivel ( $Pe_j$ ), reduciendo el posible sesgo introducido por alguno de los jueces y se obtiene mediante

$$Pe_j = \left(\frac{1}{j}\right)^j \quad (2)$$

donde  $j$  es el número de jueces expertos participantes. Posteriormente, el *CVC* se calcula de la siguiente manera (Hernández-Nieto, 2002; Hernández-Nieto, 2024; Sierra, 2023; Pedrosa, 2014).

$$CVC = CVC_i - Pe_j \quad (3)$$

Donde  $i$  representa es el número de ítem.

El *CVC* toma valores de 0 a 1. Un valor igual a 1 o superior a 0.7, con respecto a la concordancia y validez, se considera de excelente a aceptable. Un valor menor a 0.60, se considera inaceptable. En la tabla 3 se muestra los rangos e interpretaciones de los distintos valores del *CVC* de una forma más detallada (Hernández-Nieto, 2002). Hernández-Nieto (2024) recomienda mantener únicamente los ítems cuyo valor del coeficiente sea superior a 0.80.

### Tabla 3

Interpretación del cálculo del *CVC*.

Valor de CVC	Interpretación de la validez y concordancia
mayor que 0.90	Excelente
mayor que 0.80 y menor o igual a 0.90	Buena
mayor que 0.70 y menor o igual que 0.80	Aceptable
igual o mayor de 0.60 y menor o igual que 0.70	Deficiente
menor que 0.60	Inaceptable

Nota: Rangos de los valores de los CVC e interpretaciones de la validez y concordancia, de acuerdo a los rangos. Fuente: Hernández-Nieto, R. (2024). Instrumentos de Recolección de Datos en Ciencias Sociales y Ciencias Biomédicas. Universidad de Los Andes.

Para establecer el CVC se consideran los puntajes otorgados de cada juez sobre cada uno de los ítems del instrumento. El CVC total se precisa como el promedio de los CVC de cada ítem; a cada uno de ellos se le corrige por concordancia aleatoria entre jueces (Hernández-Nieto, 2024).

### 2.1. Muestra

El coeficiente de validez de contenido o CVC se obtiene por medio de cada uno de los puntajes de jueces sobre cada elemento y sobre la validez de contenido de todo el instrumento a partir del nivel de concordancia entre los jueces. Hernández-Nieto recomienda un mínimo de 3 jueces y un máximo de 5.

Se seleccionó una muestra no probabilística por conveniencia (Hernández-Sampieri, 2018) integrada por tres expertos con grado de doctorado y formación en ingeniería. Los jueces poseen una sólida trayectoria en investigación educativa y docencia en física:

- Experto 1 (U. de G.): 36 años de experiencia.
- Experto 2 (UAEH): 18 años de experiencia.
- Experto 3 (UAN): 26 años de experiencia.

La rúbrica fue digitalizada mediante herramientas de gestión de encuestas para facilitar la evaluación remota. Se solicitó a los jueces valorar cada nivel de complejidad bajo seis criterios: pertinencia, claridad conceptual, niveles cognoscitivos, formato, redacción y terminología. Se utilizó una escala tipo Likert de cinco puntos, donde 1 representa "Muy en desacuerdo" y 5 "Muy de acuerdo".

- Muy en desacuerdo (1 punto)
- En desacuerdo (2 puntos)
- Indeciso (3 puntos)
- De acuerdo (4 puntos)
- Muy de acuerdo (5 puntos)

Tras la recolección de datos, se habilitó un espacio para observaciones cualitativas que permitiera el refinamiento del instrumento. Por ejemplo, a sugerencia de uno de los jueces, se precisó el saber "Hace buen uso del equipo" integrando la cláusula "sin poner en riesgo a sus compañeros", fortaleciendo así la dimensión ética y de seguridad en la competencia de actitudes y valores.

### 3. Resultados

El análisis de las valoraciones de los jueces se realizó mediante el procesamiento de los datos en hojas de cálculo, estructurando la información por niveles de complejidad funcional (Ribes) y criterios de evaluación. Para garantizar una interpretación precisa, se aplicaron las fórmulas de Hernández-Nieto, permitiendo cuantificar el grado de consenso y validez.

Como fase inicial, se calcularon las medias de las puntuaciones otorgadas por los jueces en cada nivel, considerando un puntaje máximo de 25 puntos por bloque evaluado. A modo de ilustración, la *Tabla 3* refleja las valoraciones del Juez 3 sobre el saber “Medición de campos magnéticos en bobinas y toroides”, demostrando una consistencia elevada en los niveles de complejidad iniciales.

**Tabla 4**

*Extracto de respuestas del juez 3.*

		EVALUACIÓN JUEZ 3					TOTAL JUEZ 3
		1	2	3	4	5	
N1	Pertinencia					5	22
	Claridad conceptual				4		
	Redacción y terminología				4		
	Niveles cognoscitivos					5	
	Formato				4		
N2	Pertinencia					5	23
	Claridad conceptual				4		
	Redacción y terminología				4		
	Niveles cognoscitivos					5	
	Formato					5	

*Nota:* Ejemplo de registro de evaluación sobre el primer saber de conocimiento con respecto a los niveles de complejidad 1 y 2. *Fuente:* Elaboración propia, 2024.

Al agrupar las respuestas de cada uno de los tres jueces por niveles de la taxonomía de Ribes y promediarlas, se obtienen los resultados totales por nivel obtenidos para la evaluación de los jueces expertos, los cuales se muestran en la *Tabla 5*.

**Tabla 5**

*Promedio de los jueces expertos.*

Nivel	Juez 1	Juez 2	Juez 3
1	25	24.9	23.8
2	25	23.3	22.7
3	25	23.7	23.8
4	25	24.6	23.2
5	25	24.8	24.6

*Nota:* Promedio de los resultados de los jueces expertos para cada nivel de complejidad. *Fuente:* Elaboración propia, 2024.

Al sustituir los valores de la *Tabla 5* en las ecuaciones 1, 2 y 3 de la sección anterior se determinó el CVC para cada nivel de la taxonomía. Siguiendo el criterio de Hernández-Nieto (2024), se consideran

aceptables aquellos ítems con un valor superior a 0.80. Como se observa en la Tabla 6, todos los niveles alcanzaron valores entre 0.90963 y 0.95496, lo que sitúa al instrumento en la categoría de "excelente" en cuanto a validez de contenido.

**Tabla 6**

Valores de  $CVC$  de cada nivel de complejidad.

Nivel	$CVC_i$	$Pe_i$	$CVC_f$
1	0.98267	0.037037	0.94563
2	0.94667	0.037037	0.90963
3	0.96667	0.037037	0.92963
4	0.97067	0.037037	0.93363
5	0.99200	0.037037	0.95496

Nota: Se muestran los resultados del cálculo del  $CVC$  que corresponde a cada nivel de complejidad, donde  $CVC_i$  representa el valor de cada ítem,  $Pe_i$  al error asociado a cada nivel y  $CVC_f$  al valor final. Fuente: Elaboración propia, 2024.

Al contrastar estos hallazgos con la investigación de Sierra (2023), quien validó rúbricas para el sistema 4MAT en bachillerato, se observan coincidencias metodológicas y divergencias en los resultados. Mientras que Sierra reportó valores de  $CVC$  entre 0.8029 y 0.9096 (rangos "bueno" y "excelente"), el presente estudio alcanzó una consistencia superior, con todos sus indicadores en el rango de excelencia (superiores a 0.90).

Esta diferencia sugiere que la transposición de la taxonomía de Ribes al contexto del laboratorio de magnetismo ha permitido una definición de indicadores más precisa y unívoca para los evaluadores. Ambas investigaciones confirman que el juicio de expertos, apoyado por el coeficiente de Hernández-Nieto, es un mecanismo robusto para dotar de rigor científico a los instrumentos de evaluación que, históricamente, se han diseñado de manera empírica en el aula.

La validación por juicio de expertos confirma que la rúbrica diseñada posee un nivel de validez de contenido excelente. El alto grado de consenso ( $CVC$  máximo de 0.99200 en el Nivel 5) demuestra que los descriptores son pertinentes, claros y adecuados para medir el constructo deseado en estudiantes de ingeniería.

Este proceso no solo cumplió una función estadística, sino también cualitativa. Las recomendaciones de los jueces permitieron refinar la redacción de los saberes, asegurando que el instrumento sea comprensible y aplicable en el entorno del laboratorio.

Como prospectiva de esta investigación, se planea realizar el análisis de confiabilidad mediante el Alfa de Cronbach en una fase posterior. Asimismo, tras la aplicación piloto de la rúbrica en grupos de Ingeniería Industrial en el CUCEI, se procederá a publicar los resultados sobre el impacto real en la medición del desarrollo de competencias de los estudiantes bajo este modelo taxonómico.

## 9. Contribución de autores

Autor 1: conceptualización, investigación, metodología, redacción-borrador original. Autor 2: supervisión, visualización. Autor 3: supervisión. Autor 4: redacción-borrador original. Todos los autores: revisión y edición.

## 10. Declaración de fuentes de financiación

Este estudio se realizó con apoyo financiero del Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT).

#### 11. Declaración de uso de inteligencia artificial

Los autores declaran que, para la preparación de este trabajo, para el texto, las figuras, tablas y todas las secciones del presente manuscrito no se ha utilizado herramientas de inteligencia artificial generativa o tecnologías asistidas por IA.

#### 12. Declaración de conflicto de intereses

Los autores declaran que no existe ningún conflicto de intereses relacionado con la publicación de este artículo.

#### 4. Referencias

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2018). *Estándares para pruebas educativas y psicológicas* (M. Lieve, Trad.). American Educational Research Association. (Obra original publicada en 2014).
- Fernández, S. (2017). Evaluación y aprendizaje. *MarcoELE. Revista de Didáctica Español como Lengua Extranjera*, (24), 1–43. <https://dialnet.unirioja.es/servlet/articulo?codigo=6145807>
- García, Sergio. (2002). La Validez y la Confiabilidad en la Evaluación del Aprendizaje desde la Perspectiva Hermenéutica. *Revista de Pedagogía*, 23(67), 297-318. Recuperado en 10 de febrero de 2024, de [http://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S0798-97922002000200006&lng=es&tlng=es](http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0798-97922002000200006&lng=es&tlng=es).
- Garzón Herrera, J. F. (2023). *Aprendiendo mientras se evalúa: construcción de una rúbrica analítica centrada en la inducción magnética* [Tesis de pregrado, Universidad Pedagógica Nacional]. Repositorio Institucional UPN. <http://upnblib.pedagogica.edu.co/handle/20.500.12209/19410>
- González Canché, M. S., & Rodríguez Pérez, M. E. (2007). Evaluación de la competencia académica: un estudio de caso en la enseñanza del electromagnetismo. En Rodríguez Pérez, M. E. (Coord.), *Evaluación de los aprendizajes. Distintos enfoques* (pp. 97-126). Universidad de Guadalajara.
- Hernández–Nieto, R. (2002). *Contributions to Statistical Analysis*. Universidad de Los Andes.
- Hernández–Nieto, R. (2024). *Instrumentos de Recolección de Datos en Ciencias Sociales y Ciencias Biomédicas*. Universidad de Los Andes.
- Hernández-Sampieri, R., & Mendoza Torres, C. P. (2018). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. Editorial McGraw-Hill.
- Medina-Díaz, M. R., & Verdejo-Carrión, A. L. (2020). Validez y confiabilidad en la evaluación del aprendizaje mediante las metodologías activas. *Alteridad. Revista de Educación*, 15(2), 270-284. <https://doi.org/10.17163/alt.v15n2.2020.10>
- Pedrosa, I., Suárez-Álvarez, J., & García-Cueto, E. (2014). Evidencias sobre la validez de contenido: avances teóricos y métodos para su estimación. *Acción Psicológica*, 10(2), 3-20. <https://doi.org/10.5944/ap.10.2.11820>
- Ramírez Díaz, M. H., Olvera Aldana, M., Pérez Trejo, L., & Méndez Sánchez, A. F. (2018). *El modelo por competencias y el aprendizaje de la física*. Editorial Colofón.
- Ribes Iñesta, E., & López Valadez, F. (1985). *Teoría de la conducta: un análisis de campo y paramétrico*. Editorial Trillas.
- Ribes-Iñesta, E. (2008). Educación básica, desarrollo psicológico y planeación de competencias. *Revista Mexicana de Psicología*, 25(2), 193-207.
- Rodríguez Pérez, M. E. (2000, mayo). *Hacia una taxonomía funcional del aprendizaje útil en la elaboración de exámenes objetivos* [Ponencia]. Foro-taller de políticas y criterios para exámenes de academia, CUCEI, Universidad de Guadalajara, México.

Camelo, V., Ramírez, M., Santana, J., Rojas-Glez, A. (2026). Validez de contenido mediante juicio de expertos en rúbrica en el contexto de un laboratorio sobre temas de magnetismo

- Sánchez Sánchez, R. (2021). El tema de validez de contenido en la educación y la propuesta de Hernández-Nieto. *Latin-American Journal of Physics Education*, 15(3), 3306-1.  
<https://dialnet.unirioja.es/servlet/articulo?codigo=8358273>
- Sierra Vite, M., & Ramírez Díaz, M. H. (2023). Validación por juicio de expertos: Rúbricas de competencias en física. *Revista Internacional de Aprendizaje en Ciencia, Matemáticas y Tecnología*, 9(1), 35-46.  
<https://www.scribd.com/document/826535445/watermarked-validacion-por-juicio-de-expertos-may-06-2024-01-01-35>
- Venier, F. L., Fernández, A., Ceballos, C. D., & Esquenazi, S. (2019). Primeros resultados en el análisis de las evaluaciones finales en el área de Física y su impacto en la planificación, en el primer año de Ingeniería, en el marco de la taxonomía SOLO. *Revista de Enseñanza de la Física*, 31(1), 731-739.