

EVALUACIÓN EDUCATIVA UNIVERSITARIA EN EL LOGRO DE APRENDIZAJES POR COMPETENCIAS

Andrés Avelino Cámara Acero
Henry Giovanni Rosales Tarazona
Elver Noel Arias Hidalgo
Alember Angulo Chávez



NOVO MUNDO

Andrés Avelino Cámara Acero

Correo: acamara@unheval.edu.pe

Perfil: Doctor en Ciencias de la Educación, Magister en Gestión y Planeamiento Educativo, Licenciado en Ciencias de la Educación con especialidad en Matemática y Física.

ORCID: <https://orcid.org/0000-0002-0624-0764>

Afiliación: Universidad Nacional Hermilio Valdizán.

Henry Giovanni Rosales Tarazona

Correo: hrosales@unheval.edu.pe

Perfil: Doctor en Ciencias de la Educación, Magister en Investigación y Docencia Superior, Segunda especialidad de Educación Primaria, Licenciado en Ciencias de la Educación con especialidad en Matemática y Física.

ORCID: <https://orcid.org/0009-0006-1258-8966>

Afiliación: Universidad Nacional Hermilio Valdizán.

Elver Noel Arias Hidalgo

Correo: hrosales@unheval.edu.pe

Perfil: Maestro en Docencia en Educación Superior e Investigación, Licenciado en Ciencias de la Educación con especialidad en Matemática y Física.

ORCID: <https://orcid.org/0000-0003-2015-3445>

Afiliación: Universidad Nacional Hermilio Valdizán.

Alembor Angulo Chávez

Correo: aangulo@unheval.edu.pe

Perfil: Doctor en Ciencias de la Educación, Magister en Investigación y Docencia Superior, Licenciado en Ciencias de la Educación con especialidad en Matemática y Física.

ORCID: <https://orcid.org/0000-0003-1281-0150>

Afiliación: Universidad Nacional Hermilio Valdizán.

EVALUACIÓN EDUCATIVA UNIVERSITARIA EN EL LOGRO DE APRENDIZAJES POR COMPETENCIAS

Andrés Avelino Cámara Acero

Henry Giovanni Rosales Tarazona

Elver Noel Arias Hidalgo

Alember Angulo Chávez



Todas nuestras publicaciones son sometidas a revisión doble-ciego de pares externos (*Peer Review Double Blinded*).

Esta publicación cuenta con licencia Creative Commons Reconocimiento - NoComercial - SinObraDerivada 3.0 Unported License.



ISBN: 978-9942-694-10-2

© Andrés Avelino Camara Acero

© Henry Giovanni Rosales Tarazona

© Elver Noel Arias Hidalgo

© Alemnber Angulo Chávez

2025

DOI: <https://doi.org/10.5281/zenodo.16334772>

© Editorial Novo Mundo

Guayaquil, Ecuador.

www.editorialnovomundo.com

Queda prohibida la reproducción bajo cualquier modalidad de toda o una parte de esta obra sin autorización expresa del titular de los derechos.

Diseño de carátula y composición: Editorial Novo Mundo

Edición electrónica: Editorial Novo Mundo

Editado en Ecuador/ *Published in Ecuador*

EVALUACIÓN EDUCATIVA UNIVERSITARIA EN EL LOGRO DE APRENDIZAJES POR COMPETENCIAS

Andrés Avelino Cámara Acero

Henry Giovanni Rosales Tarazona

Elver Noel Arias Hidalgo

Alember Angulo Chávez



ÍNDICE

ÍNDICE DE TABLAS	9
ÍNDICE DE FIGURAS.....	11
INTRODUCCIÓN.....	12

CAPÍTULO I

DE LA MEDICIÓN DEL SABER A LA VALORACIÓN DE LA COMPETENCIA: UNA RUPTURA PARADIGMÁTICA EN LA EVALUACIÓN DE LAS CIENCIAS EXACTAS.....	17
1.1. La crisis del modelo evaluativo tradicional	17
1.2. El constructo de “competencia” en el dominio científico-técnico	21
1.3. La tensión epistemológica entre contenido y competencia	26
1.4. La necesidad de repensar la evaluación no como un fin, sino como un medio para el desarrollo del pensamiento científico	30

CAPÍTULO II

LA ARQUITECTURA DE LA EVALUACIÓN POR COMPETENCIAS: DISEÑO DE INSTRUMENTOS Y PROCESOS EN FÍSICA Y MATEMÁTICA	35
2.1. Principios para el diseño de una evaluación auténtica	35
2.2. Metodologías y estrategias de evaluación de competencias.....	39
2.3. Instrumentos para la valoración de procesos complejos.....	42
2.4. El rol del docente como evaluador-mediador	46

CAPÍTULO III

IMPACTO Y CONSECUENCIAS DE LA EVALUACIÓN POR COMPETENCIAS: LA DIMENSIÓN HUMANA Y SOCIAL EN LA FORMACIÓN CIENTÍFICA.....	53
3.1. Las consecuencias humanas de las prácticas evaluativas.....	53
3.2. Impacto en el sujeto aprendiz: Identidad, motivación y bienestar...	56
3.3. La evaluación por competencias y su rol en la equidad e inclusión .	60
3.4. Impacto en la cultura disciplinar y la institución universitaria	64

CAPÍTULO IV

EFFECTOS DE LA EVALUACIÓN EDUCATIVA EN APRENDIZAJES POR COMPETENCIAS EN ESTUDIANTES DE MATEMÁTICA Y FÍSICA DE LA UNHEVAL DE HUÁNUCO	69
4.1. Introducción	69
4.2. Aspectos metodológicos	73
4.3. Hallazgos	77
4.4. Discusión y conclusiones	86

REFLEXIONES FINALES.....	91
REFERENCIAS BIBLIOGRÁFICAS	97

ÍNDICE DE TABLAS

Tabla 1 Limitaciones Fundamentales del Modelo Evaluativo Tradicional en Ciencias Exactas	19
Tabla 2 Descomposición Tridimensional del Constructo de Competencia Científica	24
Tabla 3 Principios Fundamentales para el Diseño de Evaluación Auténtica en Ciencias	38
Tabla 4 Componentes de una Rúbrica para Evaluar Razonamiento Científico en la Escritura	44
Tabla 5 Dimensiones del Rol Docente como Evaluador-Mediador y su Impacto	49
Tabla 6 Distribución de Frecuencias sobre la Percepción Docente de la Evaluación Educativa (General).	78
Tabla 7 Distribución de Frecuencias de la Dimensión Finalidad y Función.	80
Tabla 8 Distribución de Frecuencias de la Dimensión Momento de Aplicación	81
Tabla 9 Distribución de Frecuencias sobre el Aprendizaje por Competencias (General)	82

Tabla 10 Distribución de Frecuencias de la Dimensión Contenido Conceptual.....	83
Tabla 11 Distribución de Frecuencias de la Dimensión Contenido Actitudinal	85

ÍNDICE DE FIGURAS

Figura 1 Niveles de la evaluación educativa.....	79
Figura 2 Niveles de la dimensión finalidad y función	80
Figura 3 Niveles de frecuencias de los aprendizajes por competencias ..	82
Figura 4 Niveles de frecuencias del significado del contenido conceptual	84
Figura 5 Niveles de frecuencias del significado del contenido actitudinal	85

INTRODUCCIÓN

La educación superior, y en particular la formación en ciencias exactas, se encuentra en una encrucijada paradigmática de profundas implicaciones para el futuro. Las demandas de una sociedad globalizada y tecnológicamente avanzada, que valora la innovación, la adaptabilidad y la resolución de problemas complejos por encima de la mera acumulación de información, han sometido a una tensión sin precedentes a los modelos pedagógicos heredados del siglo XX. En el epicentro de esta tensión se halla la evaluación educativa, un componente que, lejos de ser un mero apéndice administrativo, funciona como el sistema nervioso del currículo, determinando qué se enseña, cómo se aprende y, en última instancia, qué tipo de profesional se forma. La persistencia de prácticas evaluativas centradas en la memorización y la reproducción de conocimiento declarativo crea una disonancia crítica con la necesidad de cultivar competencias científicas sofisticadas, un desfase que este libro se propone analizar y resolver. La literatura académica reciente constata una y otra vez esta problemática, evidenciando una brecha significativa entre los perfiles de egreso que las universidades declaran perseguir y los mecanismos que utilizan para certificar su logro (Gallardo, 2020; Menzala et al., 2023).

Frente a este panorama, emerge el problema central que articula la presente obra: la profunda crisis de pertinencia del modelo evaluativo tradicional para valorar y fomentar las competencias que definen al científico y

al matemático del siglo XXI. Habilidades como la modelización de fenómenos, el pensamiento crítico, la argumentación basada en evidencia o la resolución de problemas abiertos son, por su naturaleza, refractarias a la medición mediante instrumentos estandarizados que privilegian la respuesta única y el procedimiento algorítmico (Konyeme & Alordiah, 2024; Redish & Kuo, 2015). Esta insuficiencia no es meramente técnica, sino fundamentalmente epistemológica. Revela una concepción del conocimiento científico como un producto estático y acabado, ignorando su naturaleza dinámica, contextual y eminentemente pragmática. Se analizaron los datos, se contrastaron las teorías, y la persistencia de este modelo no solo obstaculiza el desarrollo de habilidades de orden superior, sino que puede llegar a enviar un mensaje pedagógico contraproducente a los estudiantes, enseñándoles que el éxito académico reside en la conformidad y la reproducción, en lugar de en la curiosidad y la indagación.

En respuesta a este vacío, este libro defiende una tesis central: **la transición hacia la evaluación por competencias no es una mera reforma metodológica, sino una evolución epistemológica, pedagógica y ética indispensable en la formación científica, que exige una re-arquitectura sistémica de la concepción, el diseño y la función de la evaluación universitaria**. Sostenemos que superar la crisis actual requiere ir más allá de la simple sustitución de unos instrumentos por otros. Exige un cambio de paradigma que desplace el foco de la evaluación desde la medición del saber hacia la valoración de la competencia; un viraje que redefine la propia finalidad del acto evaluativo, transformándolo de una herramienta de certificación en un motor para el aprendizaje y el desarrollo del pensamiento científico. Esta reconceptualización implica una simbiosis entre contenido y competencia, donde el conocimiento disciplinar riguroso se convierte en la herramienta indispensable para la acción inteligente y reflexiva (Rott & Leuders, 2017). La presente obra se dedica a explorar en profundidad las implicaciones de esta tesis, ofreciendo no solo un diagnóstico del problema, sino también un marco de soluciones teóricamente fundamentado y empíricamente validado.

Para articular este argumento de manera coherente y progresiva, el libro se ha estructurado en cuatro capítulos interconectados, cada uno con una función lógica específica dentro del recorrido intelectual. El **Capítulo 1**, titulado *De la medición del saber a la valoración de la competencia: Una ruptura paradigmática en la evaluación de las ciencias exactas*, establece el **fundamento teórico y filosófico** de la obra. En él, se diagnostica en profundidad la crisis del modelo evaluativo tradicional, se define el constructo de “competencia” en el dominio científico-técnico para establecer un lenguaje común, se aborda la tensión epistemológica entre contenido y competencia, y se redefine la función de la evaluación como un medio para el aprendizaje. Este primer capítulo, por tanto, responde a la pregunta fundamental del “**POR QUÉ**” es necesario un cambio.

A continuación, el **Capítulo 2**, *La arquitectura de la evaluación por competencias: Diseño de instrumentos y procesos en física y matemática*, se presenta como la caja de **herramientas metodológica** de la obra. Este capítulo responde al “**CÓMO**” se diseña y se implementa una evaluación auténtica, traduciendo la teoría en práctica. Se exploran los principios rectores para el diseño de tareas evaluativas significativas, se describen las grandes estrategias metodológicas como el Aprendizaje Basado en Proyectos o la evaluación por portafolios, se detallan los instrumentos específicos para la valoración de procesos complejos —como las rúbricas analíticas—, y se analiza la transformación del rol docente hacia un perfil de evaluador-mediador. Este capítulo ofrece un andamiaje práctico para aquellos educadores e instituciones que buscan transitar hacia este nuevo paradigma.

Seguidamente, el **Capítulo 3**, *Impacto y consecuencias de la evaluación por competencias: La dimensión humana y social en la formación científica*, se dedica al **análisis de impacto**. Una vez establecido el porqué y el cómo, esta sección explora las consecuencias de este cambio paradigmático, respondiendo a la pregunta “**Y ENTONCES, ¿QUÉ PASA?**”. Se examinan los efectos de las prácticas evaluativas en la dimensión humana del aprendizaje, incluyendo el impacto en la identidad, la motivación y el

bienestar del estudiante. Asimismo, se analiza el rol de la evaluación por competencias como una posible herramienta para la equidad y la inclusión, y se exploran las repercusiones de este modelo en la cultura disciplinar y en la institución universitaria en su conjunto.

Finalmente, el **Capítulo 4**, *Efectos de la evaluación educativa en aprendizajes por competencias en estudiantes de matemática y física de la UNHEVAL de Huánuco*, funciona como la **validación empírica** de toda la obra. Este capítulo aterriza la discusión teórica, metodológica y filosófica en la realidad concreta de un estudio de caso, presentando los hallazgos de una investigación realizada en una universidad peruana. Al describir la metodología empleada, los resultados obtenidos y la discusión de los mismos, este capítulo no solo ilustra el modelo en acción, sino que proporciona una evidencia empírica que dialoga con, y enriquece, los argumentos desarrollados en las secciones precedentes. De esta manera, el libro cierra el ciclo, conectando la teoría global con la práctica local y demostrando la viabilidad y la pertinencia de la evaluación por competencias en el contexto real de la formación científica universitaria.

CAPÍTULO I

DE LA MEDICIÓN DEL SABER A LA VALORACIÓN DE LA COMPETENCIA: UNA RUPTURA PARADIGMÁTICA EN LA EVALUACIÓN DE LAS CIENCIAS EXACTAS

1.1. La crisis del modelo evaluativo tradicional

El paradigma evaluativo que ha imperado históricamente en la educación superior, y con especial arraigo en las ciencias exactas, se encuentra en una encrucijada epistemológica de profunda significación. Este modelo, fundamentado en la cuantificación del conocimiento declarativo a través de instrumentos estandarizados, enfrenta una crisis de pertinencia y validez al ser confrontado con las exigencias de una sociedad del conocimiento que demanda no solo saber, sino saber actuar (Gallardo, 2020). La premisa de que un examen escrito o una prueba de respuesta única puede capturar la complejidad del pensamiento científico se ha vuelto insostenible, revelando una disonancia fundamental entre lo que se enseña, lo que se evalúa y lo que significa ser competente en los dominios de la matemática y la física. Investigaciones recientes constatan que, incluso en programas universitarios que formalmente adoptan enfoques por competencias, persisten vestigios de un tradicionalismo evaluativo que obstaculiza el desarrollo de habilidades superiores (Castrillo & Triminio,

2024). Esta persistencia no es una mera inercia pedagógica; representa una crisis conceptual que obliga a deconstruir la propia naturaleza de la evaluación para transitar de una simple herramienta de certificación a un dispositivo catalizador del aprendizaje genuino.

Precisamente por esta disonancia, el andamiaje del modelo tradicional evidencia su incapacidad para trascender la superficie del conocimiento. Su enfoque se centra casi con exclusividad en la reproducción de información, privilegiando la memoria sobre el razonamiento y la aplicación de fórmulas sobre la modelización de fenómenos complejos. Un análisis metódico de las prácticas evaluativas en educación superior expone que el modelo hegemónico reduce los resultados del aprendizaje a “calificaciones y estadísticas”, omitiendo la valoración integral de procesos, desempeños y actitudes (Gallardo, 2020). Esta visión reduccionista es particularmente perniciosa en las ciencias exactas, donde el conocimiento no es un repositorio estático de hechos, sino un conjunto dinámico de herramientas conceptuales y procedimentales para interpretar, predecir y transformar la realidad. El estudio de Castrillo y Triminio (2024) documenta empíricamente esta problemática en carreras de Matemáticas y Física, donde la predilección docente por el examen escrito se correlaciona con elevados índices de reprobación, sugiriendo que dichos instrumentos funcionan más como filtros de conocimiento factual que como medidores de una auténtica comprensión disciplinar.

Derivado de este enfoque instrumental, la crisis se manifiesta en la inadecuación de las herramientas evaluativas para capturar las habilidades de orden superior, que constituyen el núcleo del pensamiento científico. Competencias como el pensamiento crítico, la transferencia de saberes a contextos novedosos, la toma de decisiones fundamentada y la resolución de problemas abiertos son, por su propia naturaleza, refractarias a la medición mediante pruebas estandarizadas (Gallardo, 2020). La evaluación tradicional, al atomizar el saber en unidades discretas y descontextualizadas, fracasa en valorar la capacidad del estudiante para integrar conocimientos, habilidades y actitudes en una actuación holística

frente a un desafío auténtico. Esta limitación estructural se ve reflejada en la subutilización de estrategias evaluativas de carácter experiencial—como el análisis de casos, los proyectos de investigación o las simulaciones con software especializado—que sí permitirían observar dichos desempeños complejos en acción (Castrillo & Triminio, 2024). Se configura así una paradoja: las universidades forman para la complejidad, pero evalúan para la simplicidad, generando una brecha insalvable entre el perfil de egreso deseado y las prácticas evaluativas que lo deberían construir. La siguiente tabla sintetiza las deficiencias clave de este modelo, a partir de la evidencia aportada por la literatura reciente.

Tabla 1 Limitaciones Fundamentales del Modelo Evaluativo Tradicional en Ciencias Exactas

Deficiencia Clave	Evidencia (Gallardo, 2020)	Evidencia (Castrillo & Triminio, 2024)	Implicación Epistemológica
Reduccionismo Cuantitativo	La evaluación se reduce a "calificaciones y estadísticas", perdiendo de vista el proceso de aprendizaje.	La preferencia por el examen escrito se asocia a altos índices de reprobación en áreas fundamentales.	Confunde la medición de un producto final con la valoración del desarrollo del pensamiento científico.
Fomento del Conocimiento Inerte	El modelo es incapaz de evaluar la "transferencia de conocimientos" a situaciones reales y complejas.	Se constata una subutilización de estrategias experienciales (proyectos, simulaciones) que promueven la aplicación.	Promueve un saber descontextualizado e inaplicable, contrario a la naturaleza pragmática de la ciencia.

Invisibiliza- ción de Pro- cesos Cogni- tivos	Falla en capturar desempeños com- plejos como la toma de decisiones y la resolución de pro- blemas.	Se identifican indi- cios de un tradicio- nalismo que evalúa principalmente el conocimiento de- clarativo y memo- rístico.	Ignora que la competen- cia científica reside más en el proceso de indaga- ción que en la respuesta correcta.
Desarticula- ción Holís- tica	No permite la in- tegración de cono- cimientos, habili- dades y actitudes en una evaluación integral.	La comodidad doc- cente en evaluar contenidos decla- rativos crea una tensión con la ne- cesidad de formar competencias inte- grales.	Fragmenta al sujeto aprendiz, contradiciendo la visión de un profesio- nal que actúa integrando múltiples saberes.

Fuente: Elaboración propia a partir de los hallazgos de Gallardo (2020) y Castrillo & Triminio (2024).

Esta arquitectura evaluativa, anclada en una epistemología positivista ya superada, no solo es insuficiente, sino que puede llegar a ser contraproducente para la formación científica. Al definir el éxito académico como la capacidad de replicar procedimientos y memorizar definiciones, el modelo tradicional puede desincentivar activamente la curiosidad, la creatividad y la asunción de riesgos intelectuales, que son el motor de la verdadera indagación científica (Gallardo, 2020).

La investigación de Castrillo y Triminio (2024) apunta a una “comodidad docente” en la evaluación de contenidos declarativos, una práctica que, si bien puede simplificar el proceso de calificación, perpetúa un ciclo de aprendizaje superficial. Este sistema crea un “currículo oculto” donde el estudiante aprende que la estrategia más eficiente para la aprobación no es la comprensión profunda, sino la optimización de la memorización para un formato de prueba predecible.

En consecuencia, el diagnóstico de la crisis del modelo evaluativo tradicional no es un mero ejercicio de crítica pedagógica, sino una constatación de su obsolescencia funcional y epistemológica. Este sistema, concebido para una era industrial de estandarización, resulta anacrónico en un contexto que valora la adaptabilidad, la innovación y el pensamiento complejo. La evidencia demuestra que su estructura es fundamentalmente incompatible con la valoración de lo que realmente importa en la formación de un científico o un matemático: la capacidad de usar el conocimiento como una lente para interpretar el mundo y como una herramienta para resolver sus problemas (Gallardo, 2020; Castrillo & Triminio, 2024). El modelo no está simplemente “roto”; está diseñado para un propósito distinto al que actualmente se le exige.

Por todo lo expuesto, la superación de esta crisis no pasa por ajustes menores o reformas cosméticas a los instrumentos existentes, se analizaron los datos, se contrastaron las teorías, y el resultado es inequívoco. Requiere una ruptura paradigmática que desplace el foco de la medición del saber acumulado hacia la valoración de la actuación competente. Este desplazamiento implica repensar la evaluación desde sus cimientos, cuestionando qué se valora, cómo se valora y, fundamentalmente, para qué se valora. La insuficiencia demostrada del modelo tradicional abre, por tanto, la necesidad imperiosa de explorar un nuevo constructo que pueda dar respuesta a estos interrogantes—como se advertirá en la siguiente sección—, un constructo que sitúe la competencia en el centro del proceso formativo y evaluativo.

1.2. El constructo de “competencia” en el dominio científico-técnico

Frente a la insuficiencia epistemológica del modelo tradicional, documentada en el apartado anterior, emerge la necesidad de un andamiaje conceptual alternativo que responda al vacío dejado por la evaluación sumativa. Este andamiaje se articula en torno al constructo de “competencia”, un término que trasciende la simple sinonimia con “habilidad” u “objetivo de aprendizaje” para proponer una visión integral

y dinámica del saber.

En el dominio específico de las ciencias exactas, una competencia no representa la posesión de un fragmento de conocimiento o la ejecución aislada de un procedimiento, sino la capacidad movilizadora de un individuo para orquestrar un complejo repertorio de recursos internos y externos con el fin de resolver problemas auténticos y relevantes de manera eficaz y reflexiva. Se trata de una concepción que redefine la meta educativa, desplazándola desde el “saber qué” (conocimiento declarativo) hacia el “saber actuar”, una premisa que exige la integración sinérgica de múltiples dimensiones del aprendizaje humano (Chiu & Lin, 2019; Tiruneh et al., 2017).

En respuesta a este vacío conceptual, la competencia se define como una actuación estratégica y contextualizada. A diferencia de un objetivo, que describe un resultado esperado, la competencia describe un proceso de desempeño en sí mismo. Por ejemplo, en el campo de la física o la matemática, no se trata simplemente de “resolver ecuaciones diferenciales”, sino de poseer la competencia para “modelizar un fenómeno físico mediante el lenguaje matemático apropiado, seleccionando, aplicando y validando el modelo diferencial pertinente”.

Esta última formulación implica no solo el conocimiento de las ecuaciones (el saber), sino también la destreza para aplicarlas (el saber hacer) y el juicio crítico para evaluar la idoneidad y las limitaciones del modelo (el saber ser). La investigación de Chiu y Lin (2019) sobre la competencia de modelización científica evidencia esta complejidad al definirla como la capacidad de construir, usar, comparar, revisar y comprender la naturaleza de los modelos para explicar o predecir fenómenos. Este enfoque integrador, que abarca tanto las prácticas de modelización como el metaconocimiento sobre las mismas, la posiciona como un pilar de la alfabetización científica contemporánea.

Esta visión holística del constructo se distancia radicalmente de la atomización del aprendizaje que caracteriza al modelo tradicional. Una habilidad, como la de calcular una integral o balancear una reacción química, es un componente necesario pero no suficiente de una competencia. La competencia emerge cuando el estudiante es capaz de identificar en una situación problemática no estructurada la necesidad de emplear dicha habilidad, de combinarla con otros saberes conceptuales y de ejecutarla con un propósito claro y una actitud ética y rigurosa.

El trabajo de Tiruneh et al. (2017) en el desarrollo de un test para el pensamiento crítico en física ilustra esta distinción, definiendo dicha competencia como “la capacidad de responder razonablemente a tareas de pensamiento crítico que requieren conocimiento específico del dominio”. Se constata así que la pericia científica no reside en una habilidad genérica de pensamiento crítico, sino en su manifestación situada y alimentada por un profundo conocimiento disciplinar, integrando razonamiento, análisis de argumentos y prueba de hipótesis dentro del marco conceptual de la física. Se analizaron los datos, se contrastaron las teorías, y la conclusión es que la competencia es más que la suma de sus partes; es una cualidad emergente que se manifiesta en la acción.

Para operacionalizar esta distinción fundamental, es útil descomponer el constructo en su tríada constitutiva: el saber, el saber hacer y el saber ser. El saber alude al cuerpo de conocimientos teóricos, conceptuales y factuales que el estudiante debe dominar (los principios de la termodinámica, el cálculo vectorial, etc.). El saber hacer se refiere al dominio de habilidades procedimentales y metodológicas (diseñar un experimento, programar un algoritmo, manipular instrumentos de laboratorio). Finalmente, el saber ser engloba las actitudes, los valores y las disposiciones metacognitivas (rigor intelectual, curiosidad, escepticismo organizado, conciencia de las implicaciones éticas, capacidad de autoevaluación y colaboración).

El modelo tradicional se ha enfocado casi exclusivamente en el primer componente, asumiendo erróneamente que los otros dos se desarrollarían por añadidura, una hipótesis desmentida por la práctica educativa (Chiu & Lin, 2019; Tiruneh et al., 2017). La verdadera competencia científica, por tanto, reside en la interconexión dinámica de estas tres dimensiones. La articulación de estas dimensiones se puede visualizar de manera más clara al descomponer una competencia científico-técnica paradigmática, como la resolución de problemas complejos, en el marco de esta tríada conceptual.

Tabla 2 El pH por muestra de compost

Dimensión	Definición Conceptual	Componentes Específicos en Ciencias Exactas	Referente en la Literatura
Saber (Conocimiento)	Dominio del corpus teórico, conceptual y factual de la disciplina.	-Principios y leyes fundamentales (ej. Leyes de Newton). -Modelos teóricos (ej. Modelo Estándar). -Lenguajes formales (ej. cálculo, álgebra lineal).	Conocimiento específico del dominio como prerrequisito para el pensamiento crítico (Tiruneh et al., 2017).

Saber Hacer (Habilidades)	Dominio de procedimientos, métodos y técnicas para actuar sobre el conocimiento.	-Construcción y uso de modelos para explicar/predecir. -Revisión y evaluación de modelos ante nueva evidencia. -Diseño y ejecución de procedimientos experimentales/algorítmicos.	El símbolo f representa una relación funcional. La expresión x^2 significa "el número x multiplicado por sí mismo". El gráfico es una parábola.
Saber Ser (Actitudes y Disposiciones)	Interiorización de valores, actitudes y procesos metacognitivos que guían la acción competente.	-Análisis de la incertidumbre y la probabilidad. -Razonamiento crítico y evaluación de argumentos. -Comprensión de la naturaleza, propósito y limitaciones de los modelos.	Se puede usar para modelar el área de un cuadrado de lado x . Se utiliza en física para describir la trayectoria de un proyectil.

Fuente: Elaboración propia a partir de los marcos conceptuales de Chiu & Lin (2019) y Tiruneh et al. (2017).

Esta caracterización sistémica del constructo tiene implicaciones profundas para la pedagogía y la evaluación. Al entender la competencia como un sistema integrado, se reconoce que su desarrollo no puede ser lineal ni fragmentado. Requiere de experiencias de aprendizaje que confronten al estudiante con problemas complejos, que le exijan movilizar simultáneamente sus conocimientos, sus habilidades y sus disposiciones. La evidencia psicométrica, como la obtenida en la validación del instrumento de Tiruneh et al. (2017), demuestra que es posible diseñar tareas que requieran esta integración y, por ende, evaluar la competencia

de una manera más holística y auténtica, superando la limitación de valorar únicamente el conocimiento inerte.

Por lo tanto, al definir la competencia de esta manera, se establece el vocabulario y el marco conceptual que informarán todo el libro. No se trata de una mera redefinición terminológica, sino de un cambio de foco fundamental que pone en primer plano la capacidad de actuación inteligente en contextos reales como el verdadero fin de la educación científica superior. Esta conceptualización robusta y operativizable del constructo de competencia es el pilar sobre el cual se deben edificar nuevas arquitecturas de evaluación—objeto de análisis del Capítulo 2—y es la lente a través de la cual se deben examinar las consecuencias humanas y sociales de este cambio paradigmático, que se explorarán en el Capítulo 3. Sin una comprensión clara y profunda de lo que es una competencia, cualquier intento de medirla o fomentar su desarrollo estaría condenado al fracaso.

1.3. La tensión epistemológica entre contenido y competencia

La transición hacia un paradigma evaluativo centrado en competencias, si bien responde a las limitaciones del modelo tradicional, no está exenta de controversias. De hecho, suscita una tensión epistemológica fundamental que se erige como una de las principales objeciones académicas a su implementación: la percepción de que un enfoque en la competencia podría diluir o incluso abandonar el rigor del contenido disciplinar que ha sido el pilar de la formación en ciencias exactas. Este debate, a menudo planteado como una dicotomía irreconciliable entre “saber” y “saber hacer”, se origina en una concepción del conocimiento que lo considera un corpus estático y autosuficiente.

Sin embargo, un análisis más profundo revela que esta polarización es falaz; la competencia no se opone al contenido, sino que lo resignifica, transformándolo de un fin en sí mismo a una herramienta indispensable para la acción inteligente y contextualizada. Investigaciones en la didáctica

de la física y la matemática demuestran que el verdadero dominio científico no reside en la mera posesión de un conocimiento formal, sino en la capacidad de dotar a dicho conocimiento de significado físico y funcional dentro de un contexto específico (Redish & Kuo, 2015; Rott & Leuders, 2017).

Explorando esta relación simbiótica, se constata que la competencia científica requiere, por definición, un dominio profundo y estructurado del contenido disciplinar. Un estudiante no puede modelizar competentemente un sistema dinámico sin un conocimiento robusto del cálculo diferencial, ni puede evaluar críticamente un argumento estadístico sin una comprensión sólida de la teoría de la probabilidad. La diferencia crucial radica en la teleología del aprendizaje: mientras que el modelo tradicional posiciona la acumulación de contenido como el objetivo final, el enfoque competencial lo concibe como el sustrato necesario para el desarrollo de capacidades de orden superior. El análisis metódico de Redish y Kuo (2015) ilumina esta dinámica al diferenciar la “sintaxis formal” de las expresiones matemáticas de su “significado físico”. Argumentan que, aunque las ecuaciones puedan ser idénticas en matemática y física, en esta última adquieren una carga semántica vinculada a la medición, las condiciones de aplicabilidad y las metáforas encarnadas, una resignificación que es, en sí misma, una competencia. El contenido, por tanto, no se descarta; se activa, se contextualiza y se pone al servicio de la interpretación del mundo real.

Esta resignificación del contenido encuentra sus raíces en un viraje epistemológico desde el positivismo hacia perspectivas más constructivistas y pragmáticas del conocimiento científico. El positivismo, que subyace al modelo evaluativo tradicional, concibe el conocimiento como un conjunto de verdades objetivas y universales que el estudiante debe asimilar. En contraposición, una visión constructivista o pragmática entiende el conocimiento como una construcción activa y situada, una herramienta que adquiere su valor en su capacidad para resolver problemas y dar sentido a la experiencia.

Se analizaron los datos, se contrastaron las teorías, y es evidente que desde esta última perspectiva, la competencia es la manifestación más elevada del dominio del contenido, ya que implica no solo su posesión, sino su uso estratégico y reflexivo. El estudio de Rott y Leuders (2017) sobre las creencias epistemológicas de futuros docentes de matemáticas ofrece un respaldo empírico a esta idea, mostrando que la competencia en pensamiento crítico matemático no se correlaciona con una posición específica sobre la certeza del conocimiento (positivista vs. falibilista), sino con la sofisticación de la justificación argumentativa.

Lo que esto evidencia es que el desarrollo de competencias puede ser epistemológicamente agnóstico respecto a la naturaleza última del contenido, pero exigente en cuanto a la capacidad de razonar con y sobre él. La tensión, por lo tanto, no es entre contenido y competencia, sino entre una pedagogía de la transmisión y una pedagogía de la construcción. El modelo tradicional, al evaluar la reproducción del contenido, promueve una epistemología personal en el estudiante que valora la autoridad externa y la memorización.

En cambio, un modelo que evalúa la competencia—entendida como la capacidad de usar el contenido para argumentar, modelizar y resolver problemas—fomenta una epistemología más sofisticada, en la que el estudiante se convierte en un agente activo en la construcción y validación del conocimiento. La investigación de Rott y Leuders (2017) es categórica al respecto: la justificación de la creencia (un indicador de sofisticación argumentativa) es un predictor significativo de las competencias matemáticas ($F = 11.386$; $p = 0.001$), mientras que la posición de creencia (certeza vs. incertidumbre) no lo es.

La resolución de esta aparente dicotomía, por ende, es crucial para justificar la rigurosidad de los métodos evaluativos que se propondrán en el Capítulo 2 de esta obra. Si la competencia se entendiera como una serie de habilidades genéricas y descontextualizadas, las críticas sobre la superficialidad del enfoque estarían justificadas. Sin embargo, al definirla

como una actuación situada que moviliza un conocimiento disciplinar profundo, se establece que evaluar competencias es, de hecho, una forma más exigente y auténtica de valorar el dominio del contenido.

Redish y Kuo (2015) lo expresan con claridad al sostener que en ciencia “no solo usamos matemáticas, sino que construimos significado con ellas de manera diferente a como lo hacen los matemáticos”, una habilidad interpretativa que va mucho más allá de la simple aplicación de fórmulas. La evaluación por competencias, por tanto, no se pregunta “¿qué sabe el estudiante?”, sino “¿qué es capaz de hacer el estudiante con lo que sabe?”, una interrogante que implica necesariamente la primera.

Consecuentemente, la tensión entre contenido y competencia se disuelve cuando se adopta una perspectiva epistemológica que valora el conocimiento no por su volumen, sino por su potencial de acción. Este marco filosófico permite entender que las metodologías de evaluación basadas en proyectos, estudios de caso o portafolios no son “meras actividades” que sacrifican el rigor, sino escenarios complejos diseñados precisamente para provocar y hacer visible la simbiosis entre el saber teórico y la actuación competente. Dichos métodos no ignoran el contenido; exigen su movilización, integración y aplicación en contextos que simulan los desafíos reales de la práctica científica.

En definitiva, la falsa dicotomía entre contenido y competencia es quizás el obstáculo conceptual más significativo para la modernización de la evaluación en ciencias. Superarlo requiere un profundo ejercicio de reflexión epistemológica por parte de la comunidad académica, un reconocimiento de que el fin último de la formación científica no es crear enciclopedias andantes, sino cultivar mentes capaces de pensar y actuar científicamente. Al establecer que la competencia es la expresión funcional y contextualizada de un contenido disciplinar riguroso, se sientan las bases para el siguiente paso de esta investigación: explorar la arquitectura de procesos e instrumentos que puedan valorar de manera justa y precisa esta simbiosis.

1.4. 1.4. La necesidad de repensar la evaluación no como un fin, sino como un medio para el desarrollo del pensamiento científico

La convergencia de los argumentos expuestos a lo largo de este capítulo —la crisis del modelo tradicional, la emergencia del constructo de competencia y la resolución de la tensión epistemológica entre contenido y desempeño— conduce inexorablemente a una reconceptualización funcional de la evaluación. Este cambio de paradigma trasciende la mera sustitución de instrumentos para proponer una transformación teleológica fundamental: la evaluación debe dejar de ser concebida como un evento terminal, un mero fin certificador del aprendizaje, para convertirse en un medio intrínseco y dinámico para su desarrollo. En el contexto de las ciencias exactas, esto implica que el propósito último de cualquier acto evaluativo no es simplemente asignar una calificación, sino cultivar activamente el pensamiento científico en los estudiantes. Investigaciones sistemáticas recientes demuestran que las prácticas evaluativas contemporáneas en indagación científica y educación matemática están experimentando esta precisa metamorfosis, integrando el andamiaje pedagógico y la regulación metacognitiva como componentes centrales del proceso evaluativo (Vo & Simmie, 2024; Hoang et al., 2024).

Al adoptar esta nueva funcionalidad, la evaluación se alinea de manera explícita con los principios de la evaluación formativa. Su objetivo principal ya no es juzgar, sino informar y guiar. Las tareas evaluativas, en lugar de ser obstáculos a superar, se convierten en oportunidades estructuradas para que los estudiantes pongan en práctica sus conocimientos, identifiquen sus errores conceptuales y refinen sus estrategias de resolución de problemas. La evidencia empírica respalda esta transición: una revisión sistemática de 63 estudios sobre la evaluación de la indagación científica revela que casi el 40% de las modalidades evaluativas ya son digitales e interactivas, incorporando a menudo retroalimentación formativa en tiempo real (Vo & Simmie, 2024). Este tipo de evaluación, enriquecida por técnicas como

el análisis de archivos de registro (log-file mining), permite un diagnóstico continuo del proceso de aprendizaje del estudiante, ofreciendo feedback específico sobre los nueve componentes clave de la indagación, como la formulación de hipótesis o el control de variables, transformando la evaluación en una herramienta de desarrollo cognitivo.

Este enfoque sitúa la metacognición en el epicentro del proceso de aprendizaje. Al concebir la evaluación como un medio, se fomenta que los estudiantes se conviertan en agentes activos de su propia formación, reflexionando sobre sus procesos de pensamiento, monitoreando su comprensión y autorregulando sus esfuerzos. Un análisis bibliométrico exhaustivo sobre la metacognición en la educación matemática identifica tres habilidades reguladoras fundamentales que son directamente cultivadas a través de una evaluación formativa: la planificación (seleccionar estrategias), el monitoreo (supervisar el progreso y detectar errores) y la evaluación (juzgar la calidad de las soluciones y la eficacia de las estrategias empleadas) (Hoang et al., 2024). Desde esta perspectiva, la retroalimentación que emana del proceso evaluativo no es solo una corrección externa, sino un insumo crucial para que el estudiante calibre su “precisión metacognitiva”, un factor que ha demostrado ser un predictor significativo del rendimiento matemático. Se analizaron los datos, se contrastaron las teorías, y el veredicto es claro: evaluar para la competencia es, en esencia, evaluar para la metacognición.

Además, esta redefinición funcional de la evaluación como un medio para el aprendizaje justifica y exige un cambio en la naturaleza de las tareas propuestas. Las preguntas de respuesta única o los problemas algorítmicos estandarizados, si bien pueden tener un lugar en la verificación de conocimientos básicos, son insuficientes para este nuevo propósito. Se necesitan tareas complejas, abiertas y auténticas que simulen los desafíos reales de la práctica científica. Estas tareas no solo hacen visible el pensamiento del estudiante, sino que lo provocan. Al enfrentarse a un problema que no tiene una solución única o un camino predefinido, el estudiante se ve obligado a movilizar de manera integrada

sus conocimientos conceptuales, sus habilidades procedimentales y sus disposiciones metacognitivas, es decir, a actuar competentemente. El análisis de Vo y Simmie (2024) sobre las performance assessments y las simulaciones interactivas como herramientas para evaluar la indagación científica ilustra perfectamente este tipo de tareas que, por su diseño, fusionan la evaluación con la propia experiencia de aprendizaje.

Esta transformación de la función evaluativa, de fin a medio, tiene profundas implicaciones para el rol del docente, un aspecto que, si bien se desarrollará en profundidad en secciones posteriores, merece una mención preliminar en este punto. El docente deja de ser un mero “calificador” para convertirse en un diseñador de experiencias de aprendizaje y un mediador del desarrollo cognitivo. Su labor se centra en crear los escenarios evaluativos, observar los desempeños, interpretar la evidencia y, fundamentalmente, proporcionar una retroalimentación efectiva que promueva la reflexión y la autorregulación en el estudiante. La evaluación, en este sentido, se convierte en el principal vehículo para el diálogo pedagógico entre el docente y el aprendiz, un diálogo centrado no en el juicio, sino en el crecimiento intelectual.

Por consiguiente, al concluir este capítulo introductorio, se establece una premisa fundamental que servirá de puente hacia la discusión metodológica del Capítulo 2. La necesidad de repensar la evaluación como un medio para el desarrollo del pensamiento científico no es una mera declaración de principios pedagógicos, sino una conclusión lógica que se deriva de la crisis del modelo anterior y de una comprensión más sofisticada de lo que significa ser competente en ciencia. Si el objetivo de la educación científica superior es cultivar mentes capaces de indagar, modelizar y resolver problemas de forma autónoma y reflexiva, entonces la evaluación debe ser la principal herramienta para lograrlo. Esta redefinición funcional no solo es legítima, sino que exige la exploración y el diseño de metodologías e instrumentos específicos—como los que se abordarán a continuación—que estén a la altura de esta nueva y ambiciosa misión. La pregunta ya no es si la evaluación puede ser un motor del aprendizaje; la evidencia actual

indica que debe serlo. El desafío, que se abordará en las páginas siguientes, es cómo construir esa maquinaria pedagógica de manera efectiva.

CAPÍTULO II

LA ARQUITECTURA DE LA EVALUACIÓN POR COMPETENCIAS: DISEÑO DE INSTRUMENTOS Y PROCESOS EN FÍSICA Y MATEMÁTICA

2.1. Principios para el diseño de una evaluación auténtica

Una vez establecida la necesidad de una transformación funcional de la evaluación en el capítulo precedente, este apartado se adentra en la arquitectura de dicho cambio, delineando los principios rectores que deben guiar el diseño de una evaluación por competencias genuinamente auténtica. Superar el modelo tradicional no implica una anarquía evaluativa, sino la adopción de un marco de diseño riguroso y fundamentado en criterios que garanticen la validez y la pertinencia de la valoración.

La evaluación auténtica, en esencia, busca replicar los desafíos y las formas de trabajo propias de un dominio disciplinar, exigiendo a los estudiantes no solo que demuestren lo que saben, sino cómo pueden utilizar ese saber para actuar de manera eficaz en situaciones complejas y realistas. Investigaciones recientes en educación científica han cristalizado estos requerimientos en una serie de principios operacionales que sirven

como pilares para la construcción de tareas e instrumentos evaluativos significativos (Konyeme & Alordiah, 2024; Tüysüz et al., 2023).

El primero y más fundamental de estos pilares es el principio de contextualización. Este criterio postula que las tareas evaluativas deben estar ancladas en problemas que reflejen las situaciones del mundo real que los profesionales de la física o la matemática enfrentan. En lugar de problemas abstractos y descontextualizados, la evaluación auténtica propone escenarios que requieren la aplicación del conocimiento científico para resolver un dilema práctico, analizar un caso real o diseñar una solución a un problema de ingeniería (Tüysüz et al., 2023). Konyeme y Alordiah (2024) argumentan que este principio es crucial para fomentar la transferencia de conocimientos, una de las debilidades centrales del modelo tradicional. Al situar la evaluación en un contexto verosímil, se promueve que el estudiante no solo active su conocimiento conceptual, sino que también desarrolle el juicio crítico para discernir qué conocimientos son pertinentes para una situación dada.

Directamente ligado al anterior, emerge el **principio de complejidad cognitiva**. A diferencia de la evaluación tradicional que tiende a atomizar el conocimiento en unidades discretas, la evaluación auténtica debe proponer tareas que sean inherentemente complejas y mal estructuradas. Esto significa que los problemas no deben tener una única vía de solución ni una respuesta predeterminada, exigiendo al estudiante la integración simultánea de múltiples conceptos, habilidades y procesos de pensamiento (Konyeme & Alordiah, 2024).

Un análisis metódico de Tüysüz et al. (2023) sobre el diseño de un sistema de evaluación auténtica para futuros docentes de ciencias demuestra cómo una tarea compleja —diseñar y construir un prototipo funcional— obliga a los participantes a integrar conocimientos de física, principios de diseño, argumentación y evaluación crítica, reflejando la naturaleza multidimensional de la competencia científica. La complejidad no es un fin en sí misma, sino un medio para provocar y hacer visible el pensamiento

de orden superior.

A la par de estos criterios, se encuentra el **principio de relevancia disciplinar y carácter procesual**. La autenticidad no solo reside en la similitud con el mundo real, sino también en la alineación con las prácticas epistémicas de la comunidad científica (Konyeme & Alordiah, 2024). Las tareas evaluativas deben reflejar cómo los físicos y matemáticos construyen y validan el conocimiento: formulando hipótesis, diseñando experimentos, analizando datos, construyendo argumentos basados en evidencia y comunicando resultados. Esto implica, necesariamente, que la evaluación no puede centrarse únicamente en el producto final. El carácter procesual de la evaluación auténtica exige valorar el razonamiento, las estrategias empleadas, las decisiones tomadas y la justificación de las mismas a lo largo de todo el proceso de resolución (Tüysüz et al., 2023). La evaluación se convierte así en una ventana al pensamiento del estudiante, permitiendo un diagnóstico mucho más rico y formativo que la simple corrección de una respuesta final. Se analizaron los datos, se contrastaron las teorías, y la conclusión es que evaluar el proceso es tan crucial como evaluar el resultado.

La implementación de estos principios se manifiesta en una arquitectura evaluativa que busca maximizar el potencial formativo y la transferibilidad del aprendizaje. Konyeme y Alordiah (2024) reportan que los estudiantes sometidos a regímenes de evaluación auténtica basados en estos pilares exhibieron un incremento del 34% en habilidades de pensamiento crítico y un 28% en la capacidad de transferencia de conocimientos en comparación con grupos de control bajo evaluación tradicional. Estos resultados cuantitativos sugieren que el diseño intencionado de la evaluación, guiado por estos principios, tiene un impacto directo y medible en el desarrollo de las competencias que se pretenden formar. La siguiente tabla sintetiza estos principios y su operacionalización en el diseño evaluativo.

Tabla 3 Principios Fundamentales para el Diseño de Evaluación Auténtica en Ciencias

Principio Rector	Definición Operacional	Implicación en el Diseño de Tareas	Evidencia de la Literatura
Contextualización	Las tareas se basan en problemas y escenarios realistas y pertinentes al dominio profesional.	Proponer estudios de caso, problemas de ingeniería, análisis de fenómenos físicos reales.	Aumenta la transferencia de conocimientos y la motivación intrínseca (Konyeme & Alordiah, 2024).
Complejidad Cognitiva	Las tareas son abiertas, mal estructuradas y requieren la integración de múltiples saberes y habilidades.	Diseñar proyectos de investigación, desafíos de modelización, problemas sin solución única.	Obliga a la integración de conocimiento, diseño y argumentación, reflejando la práctica científica real (Tüysüz et al., 2023).
Carácter Proceso	La valoración se centra tanto en el proceso de resolución como en el producto final.	Requerir la documentación del proceso (ej. bitácoras, borradores), presentaciones orales de justificación.	Permite valorar el razonamiento, la formulación de hipótesis y la revisión iterativa (Tüysüz et al., 2023).
Relevancia Disciplinar	Las tareas reflejan las prácticas epistémicas y los estándares de la comunidad científica.	Incluir la revisión por pares, la construcción de argumentos basados en evidencia, la comunicación de resultados.	Alinea la evaluación con las formas de trabajo y validación del conocimiento propias de la disciplina (Konyeme & Alordiah, 2024).

Fuente: Elaboración propia a partir de los marcos conceptuales de Konyeme & Alordiah (2024) y Tüysüz et al. (2023).

Finalmente, la adopción de estos principios rectores no es una mera elección metodológica, sino una declaración de intenciones pedagógicas. Al diseñar evaluaciones que son contextualizadas, complejas, procesuales y relevantes, se envía un mensaje claro a los estudiantes sobre qué es lo que realmente se valora en su formación: no la capacidad de memorizar, sino la capacidad de pensar y actuar como un científico. El estudio de caso de Tüysüz et al. (2023) es elocuente al demostrar que los 34 participantes, al ser evaluados mediante un proceso auténtico en seis fases, lograron autoevaluar con éxito sus propios diseños, evidenciando un desarrollo no solo de competencias técnicas, sino también de capacidades metacognitivas y de autorregulación.

Por tanto, estos principios constituyen el armazón conceptual sobre el cual se deben construir las metodologías y los instrumentos de evaluación que se detallarán en los siguientes apartados de este capítulo. Actúan como un filtro de calidad y un marco de referencia, asegurando que cualquier estrategia o herramienta que se proponga esté alineada con el objetivo fundamental de promover un aprendizaje profundo, significativo y transferible. La transición hacia la evaluación por competencias no es un salto al vacío, sino un camino guiado por un conjunto de principios sólidos y empíricamente validados que garantizan el rigor y la pertinencia del proceso evaluativo.

2.2. Metodologías y estrategias de evaluación de competencias

Anclados en los principios de autenticidad expuestos en la sección anterior, este apartado se adentra en la operacionalización de dichos principios a través de metodologías y estrategias evaluativas a gran escala. Si los principios actúan como el “qué” de la evaluación auténtica, las metodologías representan el “cómo” se orquestan las experiencias de aprendizaje para que la valoración de competencias complejas sea posible y significativa. A diferencia de un examen puntual, estas estrategias se conciben como procesos extendidos en el tiempo, que integran la instrucción, la práctica y la evaluación en un ciclo continuo. El Aprendizaje

Basado en Proyectos (ABP), los Estudios de Caso, la Resolución de Problemas Abiertos y la Evaluación por Portafolios emergen como las arquitecturas pedagógicas más robustas para encarnar los criterios de contextualización, complejidad y carácter procesual en la formación de físicos y matemáticos (Nicholus et al., 2023; Shidiq et al., 2024).

Una de las metodologías más potentes y documentadas en la literatura es el **Aprendizaje Basado en Proyectos (ABP)**, también conocido por sus siglas en inglés, PBL (Problem-Based Learning o Project-Based Learning). Esta estrategia sitúa un problema complejo y auténtico en el centro del proceso de aprendizaje. A los estudiantes se les presenta un desafío, a menudo interdisciplinario, que no puede ser resuelto mediante la aplicación directa de una fórmula o un algoritmo, sino que requiere un ciclo iterativo de investigación, formulación de hipótesis, aplicación de principios teóricos y síntesis de soluciones (Nicholus et al., 2023). Dentro de esta metodología, la evaluación no es un apéndice final, sino que está imbricada en cada fase del proyecto: se valora la calidad de las preguntas de investigación que formulan los estudiantes, la pertinencia de las fuentes que consultan, la coherencia de los modelos que proponen y la solidez de los argumentos con los que defienden sus soluciones.

La efectividad del ABP como estrategia evaluativa ha sido validada empíricamente. Una revisión sistemática de 45 estudios sobre la aplicación del ABP en la enseñanza de la física demostró mejoras competenciales sustanciales: un incremento promedio del 42% en habilidades de resolución de problemas complejos, del 36% en pensamiento crítico y del 48% en capacidades de modelización matemática (Nicholus et al., 2023). Estos hallazgos evidencian que el ABP no solo es una metodología de enseñanza eficaz, sino una potente estrategia de evaluación que permite observar y valorar competencias en acción. Al confrontar a los estudiantes con problemas que emulan los desafíos de la práctica profesional, el ABP encarna de manera natural los principios de contextualización y complejidad, haciendo de la evaluación una experiencia de aprendizaje en sí misma.

De forma complementaria, la **Evaluación por Portafolios** se presenta como una estrategia longitudinal que permite documentar y valorar el desarrollo de competencias a lo largo del tiempo. Un portafolio no es una mera colección de trabajos, sino una compilación intencionada y reflexiva de evidencias que demuestran el progreso del estudiante hacia metas de aprendizaje predefinidas. En el contexto de las ciencias exactas, un portafolios puede incluir una variedad de artefactos: soluciones a problemas complejos, registros de laboratorio, análisis de datos, modelos computacionales, reflexiones metacognitivas sobre el propio proceso de aprendizaje, e incluso retroalimentación recibida de pares y docentes (Shidiq et al., 2024). Esta metodología es inherentemente procesual y permite una visión holística y multidimensional del desempeño del estudiante.

La investigación de Shidiq et al. (2024), aunque realizada con futuros docentes, ofrece un modelo estructuralmente sólido y transferible a la educación superior. Al implementar una evaluación por portafolios en seis fases secuenciales (desde la orientación al problema hasta la reflexión final), el estudio constató el desarrollo de un espectro de habilidades de proceso científico, desde la formulación de preguntas (56.7% de logro) hasta la observación y la experimentación (98.7% y 87.3% respectivamente). Se analizaron los datos, se contrastaron las teorías, y se observó que el portafolios no solo sirve como un repositorio de evidencias, sino que promueve activamente la práctica reflexiva y la autoevaluación, componentes clave de la metacognición. Por su naturaleza, esta estrategia permite valorar tanto los productos del aprendizaje como los procesos cognitivos y metacognitivos que los subyacen, ofreciendo una rica fuente de datos para una evaluación auténtica.

Otras estrategias, como los **Estudios de Caso** y la **Resolución de Problemas Abiertos**, comparten esta filosofía evaluativa. Los estudios de caso presentan a los estudiantes un escenario real y detallado —el análisis de un fallo estructural en ingeniería, el estudio de una anomalía en datos astronómicos, o el modelado de un fenómeno epidemiológico—

y les exigen aplicar sus conocimientos teóricos para analizar la situación, diagnosticar las causas y proponer soluciones fundamentadas. La evaluación se centra en la calidad del análisis, la profundidad de la argumentación y la creatividad de las soluciones propuestas. Por su parte, la resolución de problemas abiertos desafía a los estudiantes con enunciados que deliberadamente carecen de datos suficientes o de un único camino correcto, obligándolos a hacer suposiciones razonables, a justificar sus elecciones metodológicas y a explorar múltiples vías de solución. Ambas estrategias encarnan el principio de complejidad cognitiva y preparan a los estudiantes para la incertidumbre y la ambigüedad inherentes a la práctica científica profesional.

En conclusión, la adopción de metodologías como el Aprendizaje Basado en Proyectos y la Evaluación por Portafolios representa la aplicación a gran escala de los principios de la evaluación auténtica. Estas estrategias transforman la evaluación de un evento aislado en un proceso continuo y formativo, donde el acto de ser evaluado es indistinguible del acto de aprender. Al sumergir a los estudiantes en desafíos complejos y realistas, estas metodologías no solo permiten una valoración más válida y fiable de las competencias científicas, sino que activamente las construyen. La elección de una u otra estrategia dependerá de los objetivos de aprendizaje específicos, pero todas comparten un núcleo común: el compromiso de hacer de la evaluación un motor para el desarrollo del pensamiento científico. La implementación exitosa de estas metodologías, sin embargo, depende de la existencia de herramientas de valoración adecuadas, un tema que se abordará en la siguiente sección.

2.3. Instrumentos para la valoración de procesos complejos

La implementación efectiva de las metodologías evaluativas descritas previamente, como el Aprendizaje Basado en Proyectos o la evaluación por portafolios, depende de manera crucial de la disponibilidad de instrumentos de valoración adecuados. Si las metodologías proporcionan el marco macro para la evaluación, los instrumentos

constituyen las herramientas de micro-nivel que permiten traducir los desempeños complejos de los estudiantes en juicios evaluativos rigurosos, fiables y transparentes. La valoración de procesos cognitivos sofisticados —como el razonamiento científico, el pensamiento crítico o la capacidad de modelización— no puede depender de la apreciación subjetiva del evaluador. Requiere de instrumentos estructurados que descompongan estas competencias complejas en criterios observables y que definan con claridad los diferentes niveles de desempeño. Las rúbricas analíticas, las escalas de observación y los diarios reflexivos emergen como los dispositivos más idóneos para esta tarea (Reynders et al., 2020; Timmerman et al., 2011).

Entre estos instrumentos, la rúbrica analítica se ha consolidado como la herramienta por excelencia para la evaluación de competencias. A diferencia de una simple lista de cotejo, que solo constata la presencia o ausencia de un elemento, una rúbrica analítica desglosa una competencia en varias dimensiones o criterios, y para cada criterio, describe de forma explícita cómo se manifiesta el desempeño en diferentes niveles de calidad (por ejemplo, de “novato” a “experto”). Este diseño ofrece una doble ventaja: por un lado, proporciona a los estudiantes una guía clara sobre las expectativas de la tarea, fomentando la autorregulación; por otro, dota al evaluador de un marco compartido que aumenta significativamente la objetividad y la consistencia de la calificación. Investigaciones en el ámbito STEM han demostrado que es posible desarrollar rúbricas con altos niveles de fiabilidad inter-evaluador para valorar procesos como el pensamiento crítico y el procesamiento de información (Reynders et al., 2020).

El proceso de diseño y validación de una rúbrica robusta es, en sí mismo, un ejercicio de investigación riguroso. El trabajo de Reynders et al. (2020) ilustra un modelo de validación multi-etapa que incluye desde la revisión de literatura y la consulta con expertos hasta la implementación piloto y el análisis psicométrico. La rúbrica resultante para evaluar el pensamiento crítico, por ejemplo, no se limita a criterios genéricos, sino

que operacionaliza la competencia en dimensiones observables como el “análisis de evidencias y patrones” o la “formulación de argumentos”, cada una con descriptores detallados para cuatro niveles de desempeño. El análisis de generalización de esta rúbrica arrojó un coeficiente de fiabilidad de $g = 0.85$, un valor que no solo supera los estándares habituales en instrumentos educativos ($g > 0.80$), sino que excede la fiabilidad típica de la revisión por pares profesional ($r \approx 0.30$). Estos datos evidencian que, con un diseño cuidadoso, es posible valorar procesos cognitivos complejos con un alto grado de objetividad.

De manera similar, el estudio de Timmerman et al. (2011) presenta el desarrollo de una rúbrica “universal” para evaluar el razonamiento científico a través de la escritura, un vehículo de evaluación particularmente relevante en física y matemática. Este instrumento descompone la competencia en 15 criterios específicos distribuidos a lo largo de las secciones de un informe científico (introducción, hipótesis, métodos, resultados, discusión), como se detalla en la siguiente tabla.

Tabla 4 Componentes de una Rúbrica para Evaluar Razonamiento Científico en la Escritura

Sección del Informe	Competencias Específicas Evaluadas
Introducción	- Establecimiento del contexto y los antecedentes relevantes. - Articulación clara de la pregunta de investigación.
Hipótesis	- Comprobabilidad y falsabilidad de la hipótesis. - Mérito científico y fundamentación teórica de la hipótesis.

Métodos	- Adecuación y descripción de los controles experimentales. - Potencial de replicación del diseño experimental.
Resultados	- Selección y presentación adecuada de datos (gráficos, tablas). - Realización de análisis estadísticos pertinentes y correctos.
Discusión	- Formulación de conclusiones basadas estrictamente en la evidencia. - Consideración de explicaciones alternativas y limitaciones del diseño.
Integración	- Uso y citación correcta de literatura científica primaria.

Fuente: Adaptado de los 15 criterios de la rúbrica desarrollada por Timmerman et al. (2011).

Este tipo de instrumento demuestra cómo la evaluación de la escritura científica, mediada por una rúbrica analítica, puede convertirse en una ventana privilegiada para observar el razonamiento del estudiante. Se analizaron los datos, se contrastaron las teorías, y la conclusión es que, al evaluar cómo un estudiante argumenta, justifica sus métodos y discute sus resultados, se está valorando directamente su competencia científica.

La fiabilidad de esta rúbrica también resultó ser excepcionalmente alta ($g = 0.85$), aunque el análisis reveló una variabilidad interesante en la fiabilidad de los criterios individuales (desde $g = 0.20$ hasta $g = 0.94$), lo que sugiere que algunas sub-competencias son intrínsecamente más difíciles de evaluar de manera objetiva que otras (Timmerman et al., 2011).

Además de las rúbricas, otros instrumentos como las escalas de observación y los diarios reflexivos complementan el arsenal evaluativo. Las escalas de observación son particularmente útiles para valorar desempeños en tiempo real, como el trabajo en un laboratorio o la presentación oral de un proyecto. Permiten al evaluador registrar la frecuencia o la calidad de comportamientos específicos (por ejemplo, “colabora eficazmente con sus pares” o “manipula el equipo de laboratorio con seguridad”).

Por su parte, los diarios reflexivos son instrumentos potentes para acceder a la dimensión metacognitiva del aprendizaje. Al solicitar a los estudiantes que documenten periódicamente sus reflexiones sobre sus propios procesos de aprendizaje —qué dificultades encontraron, qué estrategias utilizaron, qué cambiarían en el futuro—, se obtiene una evidencia invaluable sobre su capacidad de autorregulación.

En síntesis, la valoración de los procesos complejos que se despliegan en las metodologías de evaluación por competencias requiere de un conjunto de instrumentos específicos, diseñados con rigor y validados empíricamente. Las rúbricas analíticas, como demuestran los trabajos de Reynders et al. (2020) y Timmerman et al. (2011), se erigen como la herramienta central para esta tarea, permitiendo una evaluación objetiva, fiable y transparente de desempeños que antes se consideraban “inmesurables”. La existencia y el uso de estos instrumentos son la condición de posibilidad para que la evaluación auténtica no sea una aspiración idealista, sino una práctica académica rigurosa y efectiva. La calidad de estos instrumentos y la pericia en su aplicación, sin embargo, dependen críticamente del actor principal en este proceso: el docente, cuyo rol transformado será el objeto de análisis de la siguiente sección.

2.4. El rol del docente como evaluador-mediador

La arquitectura de una evaluación por competencias, con sus principios, metodologías e instrumentos, conforma un andamiaje técnico sofisticado. No obstante, la efectividad de esta maquinaria pedagógica no

reside en su diseño abstracto, sino en la actuación de quien la opera: el docente. El cambio hacia una evaluación auténtica y formativa precipita una transformación radical en el rol del profesorado, desplazándolo desde la posición de un mero “calificador” o certificador de conocimientos hacia la de un “evaluador-mediador”.

En este nuevo paradigma, el docente se convierte en un facilitador del aprendizaje, un diagnosticador de procesos cognitivos y un coach que guía al estudiante hacia la autonomía intelectual. Este apartado se centra en humanizar el contenido técnico del capítulo, analizando las competencias y prácticas específicas que definen este rol mediador, cuya influencia se ha demostrado determinante para el desarrollo de competencias científicas en la educación superior (Bellido et al., 2024; Guasch et al., 2010).

Este nuevo rol exige, en primer lugar, una competencia fundamental en el **diseño de situaciones evaluativas complejas**. El docente-mediador no solo aplica instrumentos, sino que diseña los escenarios de aprendizaje que los hacen necesarios y significativos. Esto implica la capacidad de crear problemas auténticos, proyectos desafiantes y casos de estudio relevantes que sirvan como catalizadores para el desarrollo de competencias (Guasch et al., 2010).

Más allá del diseño, la actuación mediadora se manifiesta en la capacidad de **observación y diagnóstico**. Durante la ejecución de una tarea compleja, el docente debe ser un observador agudo, capaz de inferir los procesos de pensamiento del estudiante a partir de sus acciones, sus preguntas y sus errores. Esta observación no busca el juicio, sino la comprensión: ¿dónde se encuentra el obstáculo epistemológico del estudiante? ¿Qué estrategia errónea está aplicando? ¿Qué conocimiento previo está fallando en activar?

A partir de este diagnóstico, la función mediadora se materializa en la provisión de una **retroalimentación efectiva**, una de las prácticas con mayor impacto en el aprendizaje. La retroalimentación en un contexto de

evaluación por competencias dista mucho de ser una simple corrección. Debe ser oportuna, específica, comprensible y orientada a la acción, enfocándose más en el proceso que en el resultado final (Bellido et al., 2024).

El docente-mediador utiliza técnicas como el **diálogo socrático** y la **formulación de preguntas potentes** para guiar al estudiante a descubrir sus propios errores y a construir sus propias soluciones. En lugar de decir “esto está mal”, pregunta “¿qué evidencia sustenta esa conclusión?” o “¿has considerado alguna explicación alternativa?”. Este enfoque no solo corrige, sino que modela el pensamiento crítico y fomenta la metacognición.

La investigación de Bellido et al. (2024) proporciona evidencia cuantitativa contundente sobre el impacto de este rol. Su estudio con 340 estudiantes universitarios de ciencias reveló que las prácticas de evaluación formativa mediadas por el docente explican el 70.8% de la variabilidad en el desarrollo de la competencia científica general ($R^2 = 0.708$). Este hallazgo es de una importancia capital, pues constata científicamente que la acción del docente como evaluador-mediador no es un factor accesorio, sino el principal predictor del éxito en el desarrollo de competencias. Dicho análisis desglosa este rol en prácticas observables, como se muestra en la siguiente tabla.

Tabla 5 Dimensiones del Rol Docente como Evaluador-Mediador y su Impacto

Dimensión de la Práctica Docente	Descripción Operacional	Impacto en Competencias (R ²)
Clarificación de Criterios	Establece y comunica de forma explícita las expectativas y criterios de evaluación.	-
Retroalimentación Oportuna	Proporciona feedback específico y a tiempo para que pueda ser utilizado en el aprendizaje.	-
Fomento de la Autoevaluación	Provee estrategias y herramientas para que el estudiante evalúe su propio progreso.	-
Facilitación de la Co-evaluación	Organiza y guía procesos de evaluación entre pares, promoviendo el aprendizaje colaborativo.	-
Ajuste de la Enseñanza	Utiliza la información de la evaluación para adaptar y modificar sus propias prácticas pedagógicas.	-
Participación Activa	Implica al estudiante como un agente activo en su propio proceso de evaluación.	-
Competencia Global	Efecto combinado de todas las dimensiones.	0.708
Investigación Científica	Competencia para indagar usando métodos científicos.	0.437
Diseño y Tecnología	Competencia para diseñar y construir soluciones tecnológicas.	0.521

Fuente: Elaborado a partir de los hallazgos de Bellido et al. (2024).

Además, el rol de evaluador-mediador implica la competencia para gestionar procesos evaluativos complejos que a menudo involucran la co-evaluación y la autoevaluación. Facilitar una sesión de evaluación entre pares requiere que el docente no solo provea una rúbrica clara, sino que también enseñe a los estudiantes a dar y recibir retroalimentación constructiva, mediando en las posibles discrepancias y asegurando que el proceso sea justo y formativo (Guasch et al., 2010). Se analizaron los datos, se contrastaron las teorías, y es manifiesto que esta gestión va más allá de lo técnico; requiere una mediación entre las dimensiones sociales y cognitivas del aprendizaje.

La transición hacia este rol no es trivial y exige un desarrollo profesional específico. El estudio de Guasch et al. (2010) sobre la formación de docentes universitarios para la enseñanza mediada por tecnología identifica un progreso competencial que va desde la conceptualización de nuevas metodologías hasta el diseño aplicado de actividades evaluativas auténticas. Este proceso evidencia que las competencias de mediación evaluativa pueden y deben ser aprendidas y practicadas. La formación debe enfocarse en dotar al profesorado de las herramientas conceptuales y prácticas para liderar estos nuevos procesos, incluyendo la capacidad de sistematizar y sintetizar los aprendizajes que emergen de las discusiones formativas.

En conclusión, el rol del docente como evaluador-mediador es el elemento que da vida y sentido a la arquitectura técnica de la evaluación por competencias. Los principios de autenticidad, las metodologías como el ABP y los instrumentos como las rúbricas son inertes sin la actuación de un profesional capaz de diseñar, observar, diagnosticar, retroalimentar y gestionar el complejo proceso del aprendizaje. La evidencia empírica confirma que esta actuación mediadora es el factor más determinante para el desarrollo de competencias científicas. Al humanizar el proceso evaluativo y centrarlo en el diálogo y el crecimiento, el docente no solo cumple una función técnica, sino que también sienta las bases para el bienestar y la motivación del estudiante, una dimensión humana y social

que será el foco central del siguiente capítulo.

CAPÍTULO III

IMPACTO Y CONSECUENCIAS DE LA EVALUACIÓN POR COMPETENCIAS: LA DIMENSIÓN HUMANA Y SOCIAL EN LA FORMACIÓN CIENTÍFICA

3.1. Las consecuencias humanas de las prácticas evaluativas

Una vez delineada la arquitectura técnica de la evaluación por competencias, resulta imperativo dirigir la mirada hacia sus efectos, desplazando el foco del “cómo” al “y entonces, qué”. Este capítulo se adentra en la dimensión humana y social de la evaluación, partiendo de una premisa fundamental: toda práctica evaluativa, independientemente de su diseño —sea este tradicional o competencial—, trasciende su función meramente técnica de medición y genera profundas consecuencias en la vida de los individuos y en la cultura de las instituciones.

La evaluación no es un acto neutro; es un evento cargado de significado que moldea la identidad, la motivación, el bienestar y las oportunidades de los estudiantes. Investigaciones recientes han comenzado a documentar de manera sistemática el vasto alcance de estas consecuencias, evidenciando

que los procedimientos de evaluación constituyen una fuente primaria de estrés académico con repercusiones significativas en la salud mental y física de los universitarios (Jones et al., 2021; Pérez-Jorge et al., 2025).

Este análisis panorámico revela que las prácticas evaluativas operan como potentes mensajes sobre lo que la institución valora, lo que define el éxito y, en última instancia, lo que significa ser un “buen estudiante”. Cuando los sistemas de evaluación se perciben como amenazantes, punitivos o injustos, pueden desencadenar una cascada de efectos psicológicos negativos.

Jones et al. (2021) identifican una tensión crítica entre la evaluación como un “desafío” estimulante y la evaluación como una “amenaza” paralizante. Cuando los estudiantes internalizan el fracaso académico no como un evento aislado y formativo (“este trabajo no fue lo suficientemente bueno”), sino como un juicio sobre su valía personal (“no soy lo suficientemente bueno”), el impacto en su bienestar puede ser devastador. Este fenómeno, exacerbado por sistemas de calificación jerárquicos y normativos, se ha asociado con un incremento de la ansiedad, la depresión, el agotamiento e incluso las ideas suicidas en la población estudiantil (Pérez-Jorge et al., 2025).

Profundizando en esta línea, la investigación contemporánea ha establecido un vínculo robusto entre el estrés académico y una variedad de problemas de salud. El análisis sistemático de Pérez-Jorge et al. (2025) confirma que la presión asociada a los periodos de evaluación no solo afecta el rendimiento académico, sino que se manifiesta en alteraciones del sueño, pérdida de apetito, inactividad física y un deterioro general de la salud.

Estas consecuencias humanas no son un mero efecto secundario del rigor académico; son el resultado directo de sistemas evaluativos que a menudo priorizan la competición y la clasificación sobre el aprendizaje y el desarrollo personal. La comprensión de estas dinámicas es crucial,

pues establece la base para argumentar que un cambio hacia la evaluación por competencias no es solo una mejora pedagógica, sino también una intervención necesaria para promover entornos de aprendizaje más saludables y sostenibles.

En este contexto, este capítulo explorará las múltiples facetas del impacto evaluativo, funcionando como un paraguas temático para las secciones subsiguientes. En primer lugar, se analizará con detenimiento el **impacto en el sujeto aprendiz**, examinando cómo las diferentes prácticas evaluativas pueden construir o erosionar la identidad académica, fomentar la motivación intrínseca o extrínseca, y promover o deteriorar el bienestar psicológico. Se analizaron los datos, se contrastaron las teorías, y la evidencia sugiere que los sistemas basados en criterios y en la retroalimentación formativa, como los propuestos en el capítulo anterior, tienen el potencial de mitigar muchas de las amenazas psicológicas asociadas a la evaluación tradicional (Jones et al., 2021).

A continuación, la perspectiva se ampliará del individuo al colectivo, abordando el rol de la **evaluación en la equidad y la inclusión**. Se investigará si la evaluación por competencias, al valorar una gama más amplia de desempeños y ofrecer múltiples vías para demostrar el aprendizaje, puede constituir una herramienta para la justicia social en la educación superior. No obstante, también se explorarán los riesgos de que, si no se implementa con cuidado, pueda generar nuevas formas de inequidad, un debate complejo que requiere un análisis matizado de las tensiones entre la estandarización y la personalización en la evaluación (Jones et al., 2021).

Finalmente, el análisis se elevará al nivel macro para examinar el **impacto en la cultura disciplinar y la institución universitaria**. Se indagará sobre las resistencias y las transformaciones que la adopción de un nuevo paradigma evaluativo genera en los departamentos académicos y en las políticas institucionales. La transición desde una cultura de la calificación hacia una cultura de la retroalimentación no es un simple cambio de

procedimiento, sino una profunda reconfiguración de los valores y las prácticas que definen a una comunidad académica.

En resumen, este capítulo se propone desvelar las capas de consecuencias humanas, sociales e institucionales que se derivan de las prácticas evaluativas. Al establecer este panorama, no se busca repetir la crítica al modelo tradicional ya realizada, sino trascenderla, enfocándose en cómo cualquier sistema de evaluación, por su mera existencia, se convierte en una fuerza poderosa que configura la experiencia universitaria en su totalidad. Comprender estas consecuencias es el primer paso para diseñar sistemas de evaluación que no solo sean técnicamente sólidos y epistemológicamente coherentes, sino también humanamente responsables.

3.2. Impacto en el sujeto aprendiz: Identidad, motivación y bienestar

Profundizando en las consecuencias humanas esbozadas en la sección anterior, el análisis debe centrarse ahora en el nexo causal más directo y significativo: el impacto de las prácticas evaluativas en la psique del sujeto aprendiz. La evaluación, lejos de ser un mero trámite administrativo, funciona como un espejo en el que el estudiante construye su propia imagen como aprendiz en un dominio específico. Particularmente en las ciencias exactas, donde a menudo persisten narrativas culturales sobre un “talento” innato, la forma en que se evalúa puede solidificar o deconstruir la identidad del estudiante como “alguien capaz en ciencias”.

La evaluación tradicional, con su énfasis en exámenes de alto riesgo y calificaciones sumativas, tiende a fomentar una “mentalidad fija”, una creencia implícita de que la habilidad es una entidad estática e inmutable. En contraste, la evaluación por competencias, al centrarse en el proceso, el desarrollo y la retroalimentación formativa, está inherentemente alineada con una “mentalidad de crecimiento”, la convicción de que la pericia se construye a través del esfuerzo y la estrategia. Esta distinción no es una sutileza pedagógica; tiene efectos directos y medibles en la motivación,

las emociones y el bienestar del estudiante, como lo demuestran recientes investigaciones experimentales en el ámbito de la física universitaria (Heyder & Pegels, 2025; Nkopodi et al., 2024).

Esta influencia en la identidad no es meramente teórica; se ha demostrado experimentalmente su capacidad para modular la experiencia de aprendizaje del universitario. El estudio de Heyder y Pegels (2025) es particularmente elocuente al respecto. En un experimento con 130 estudiantes universitarios, la exposición a un profesor de física ficticio que manifestaba una mentalidad fija (por ejemplo, “o se entiende la física o no se entiende”) fue suficiente para generar en los estudiantes una cascada de consecuencias psicológicas negativas en comparación con aquellos expuestos a un profesor con mentalidad de crecimiento.

Específicamente, los estudiantes en la condición de mentalidad fija reportaron anticipar emociones positivas significativamente menores, emociones negativas más fuertes, un mayor nivel de estrés, y, de manera crucial, una menor disposición a buscar ayuda y expectativas de éxito más bajas. Este último punto es fundamental: un sistema evaluativo que envía señales de que la habilidad es fija disuade activamente al estudiante de buscar el apoyo necesario para superar las dificultades, creando un ciclo de fracaso auto-cumplido y consolidando una identidad de “no ser bueno para la física”, lo que deteriora su capacidad para persistir frente al desafío.

Frente a este panorama, la evaluación por competencias propone una ruta alternativa que impacta directamente en la naturaleza de la motivación del estudiante. Al desplazar el foco del juicio final hacia el proceso de desarrollo, y al utilizar tareas contextualizadas y relevantes como vehículos de aprendizaje (como se discutió en la sección 2.1), la evaluación por competencias tiene el potencial de transformar la motivación extrínseca

(aprender para la nota) en motivación intrínseca (aprender para resolver el problema).

Cuando un estudiante se enfrenta a un desafío auténtico que percibe como significativo, el motor de su esfuerzo deja de ser el miedo al castigo (una mala calificación) y pasa a ser la satisfacción intelectual de la indagación y el descubrimiento. Este cambio es vital para el aprendizaje profundo y para la formación de una identidad académica positiva y resiliente. La motivación intrínseca no solo mejora el rendimiento, sino que también protege al estudiante contra el agotamiento, ya que el aprendizaje se experimenta como una actividad gratificante y no como una carga impuesta externamente. Se analizaron los datos, se contrastaron las teorías, y es evidente que el diseño evaluativo es un arquitecto clave de la arquitectura motivacional del estudiante.

Más allá de la construcción de la identidad y el fomento de la motivación, el impacto de la evaluación se extiende a la esfera más visceral del bienestar emocional del estudiante. Las disciplinas STEM, y en particular la física y la matemática, son reconocidas como áreas que pueden generar altos niveles de ansiedad, a menudo exacerbada por prácticas evaluativas tradicionales de alto impacto (Nkopodi et al., 2024).

Los exámenes de una única oportunidad, donde un semestre de trabajo se condensa en unas pocas horas, crean un entorno de alta presión que puede ser perjudicial para la salud mental, como se introdujo en la sección 3.1. La evaluación por competencias, al ser continua, procesual y diversificada, distribuye el “riesgo” evaluativo a lo largo del tiempo y a través de múltiples tipos de evidencia. Esta diversificación no solo ofrece una imagen más válida de la capacidad del estudiante, sino que también reduce la ansiedad paralizante asociada a los eventos evaluativos únicos. La incorporación de retroalimentación formativa constante, por ejemplo, transforma la incertidumbre en claridad y el miedo al error en una oportunidad de aprendizaje, contribuyendo directamente a un entorno psicológicamente más seguro.

De hecho, la investigación ha comenzado a identificar criterios específicos para el diseño de entornos de aprendizaje en STEM que promuevan activamente el bienestar estudiantil. El análisis de Nkopodi et al. (2024), basado en una síntesis de 44 artículos, propone un enfoque holístico que incluye desde la integración de prácticas de mindfulness hasta el uso de tecnologías que fomenten la confianza. Entre sus hallazgos, destaca el rol de la retroalimentación inmediata, facilitada por ejemplo a través de pizarras interactivas o plataformas digitales, como un factor que aumenta la autoeficacia académica y reduce la ansiedad.

Cuando un estudiante recibe confirmación o corrección al instante, el ciclo de duda y temor se interrumpe, permitiéndole avanzar con mayor seguridad. Este principio es fundamental en la evaluación por competencias, donde el feedback no es un post-mortem, sino una guía en tiempo real. La implementación de programas de apoyo entre pares, como la co-evaluación estructurada, también contribuye a este bienestar al crear una comunidad de aprendizaje y reducir el aislamiento que a menudo sienten los estudiantes en disciplinas técnicamente exigentes.

Al entrelazar estos hilos, emerge una visión sistémica del impacto de la evaluación en el individuo. La identidad, la motivación y el bienestar no son constructos aislados, sino que están íntimamente interconectados en un sistema de retroalimentación dinámica. Una práctica evaluativa que promueve una mentalidad de crecimiento (Heyder & Pegels, 2025) fortalece el autoconcepto del estudiante como alguien capaz de aprender y mejorar.

Esta identidad positiva alimenta la motivación intrínseca, lo que conduce a un mayor compromiso y a la búsqueda proactiva de ayuda. El éxito derivado de este compromiso, a su vez, reduce la ansiedad y mejora el bienestar emocional, lo que refuerza la identidad positiva inicial, creando así un ciclo virtuoso de aprendizaje y desarrollo personal. Por el contrario, un sistema evaluativo que fomenta una mentalidad fija puede desencadenar un ciclo vicioso: el miedo al fracaso inhibe la motivación,

lo que lleva a un menor esfuerzo y a evitar los desafíos, resultando en un desempeño deficiente que confirma la creencia negativa inicial y deteriora el bienestar.

En conclusión, la elección entre un paradigma evaluativo tradicional y uno basado en competencias no es una decisión meramente técnica, sino una que tiene profundas implicaciones para la salud psicológica y el desarrollo integral del estudiante. Al centrarse en el proceso, ofrecer retroalimentación formativa, utilizar tareas auténticas y promover una mentalidad de crecimiento, la evaluación por competencias se posiciona como un enfoque que no solo es más válido y fiable para medir el aprendizaje complejo, sino que también es fundamentalmente más humano.

Este enfoque crea las condiciones para que el estudiante no solo aprenda física o matemáticas, sino que se desarrolle como una persona segura, motivada y resiliente, capaz de enfrentar los desafíos de su disciplina con confianza y curiosidad. Si bien este análisis se ha centrado en la experiencia psicológica individual, es crucial reconocer que estas dinámicas no ocurren en un vacío social. El impacto de la evaluación en la identidad y el bienestar tiene implicaciones directas en la equidad y la inclusión, una dimensión que se explorará en la siguiente sección.

3.3. La evaluación por competencias y su rol en la equidad e inclusión

Tras examinar el impacto de la evaluación en la psique individual del estudiante, la perspectiva debe ahora ampliarse para abarcar su dimensión social y sistémica. La evaluación, como uno de los mecanismos más potentes de poder dentro de la academia, no solo configura la experiencia individual, sino que también actúa como un clasificador social que puede perpetuar o mitigar las inequidades existentes. El foco de este análisis se desplaza, por tanto, hacia una pregunta de crucial importancia para la justicia social en la educación superior: ¿puede la evaluación por competencias funcionar como una herramienta para promover la equidad

y la inclusión?

La premisa fundamental es que, al alejarse de los modelos de evaluación tradicionales —a menudo basados en pruebas estandarizadas de alto riesgo que favorecen ciertos capitales culturales y estilos de aprendizaje—, la evaluación por competencias, con su énfasis en múltiples formas de evidencia y trayectorias de aprendizaje personalizadas, posee el potencial inherente de crear un campo de juego más equitativo. Sin embargo, este potencial no es automático y su realización depende de un diseño intencional y reflexivo que evite la creación de nuevas barreras sistémicas (Richard et al., 2022; Marion et al., 2020).

Esta promesa de equidad encuentra un sólido respaldo empírico en estudios que han analizado los efectos diferenciales de la implementación de sistemas de calificación basados en competencias. Un análisis cuasi-experimental de gran relevancia, realizado en un curso introductorio de física, constató que la transición de un sistema de calificación tradicional a uno basado en el dominio de competencias (el sistema ALPaCA) no solo mejoró el rendimiento general de los estudiantes, sino que lo hizo de manera desproporcionadamente beneficiosa para grupos históricamente subrepresentados en STEM.

Específicamente, el estudio evidenció que, si bien todos los estudiantes mejoraron, las ganancias más significativas se observaron en mujeres y en estudiantes de primera generación (Richard et al., 2022). Las tasas de DFW (calificaciones de D, F o abandono) se redujeron de manera estadísticamente significativa para ambos grupos ($p=0.017$ para mujeres; $p=0.05$ para estudiantes de primera generación). Este hallazgo es de suma importancia, pues sugiere que la evaluación por competencias puede actuar como un mecanismo de nivelación, mitigando las barreras sistémicas que a menudo enfrentan estos colectivos.

El mecanismo subyacente a este efecto equitativo parece residir en la deconstrucción de las barreras que impone la evaluación tradicional. El estudio de Richard et al. (2022) también reveló que la implementación del sistema ALPaCA redujo la dependencia de los puntajes de pruebas estandarizadas previas (como el ACT Math) como predictores del éxito en el curso. La evaluación tradicional, con su formato de examen cronometrado y de respuesta única, tiende a valorar la velocidad de procesamiento y la familiaridad con un tipo específico de prueba, habilidades que pueden estar más correlacionadas con el capital socioeconómico y educativo previo que con la capacidad intrínseca para aprender física.

En contraste, un sistema de evaluación por competencias, al ofrecer múltiples oportunidades para demostrar el dominio y al valorar una gama más amplia de desempeños a través de diversas tareas, permite que estudiantes con diferentes estilos de aprendizaje y ritmos de progreso puedan evidenciar su comprensión y habilidad de manera más auténtica y justa. Se analizaron los datos, se contrastaron las teorías, y la conclusión es que este enfoque desplaza el foco desde el “potencial” predicho por pruebas estandarizadas hacia el “desempeño” demostrado a través del trabajo sostenido.

Para materializar este potencial, es indispensable que los sistemas de evaluación por competencias se diseñen con la equidad como un principio rector explícito. No se trata simplemente de cambiar las tareas, sino de construir un “sistema de evaluación balanceado” que integre múltiples fuentes de evidencia para informar tanto el aprendizaje del estudiante como la mejora continua del propio sistema (Marion et al., 2020).

Un sistema de este tipo utiliza una combinación de evaluaciones formativas de bajo riesgo, tareas de desempeño a nivel de unidad y evaluaciones sumativas que permiten una visión holística y multifacética del progreso de cada estudiante. Esta multiplicidad de evidencias es intrínsecamente más equitativa, ya que reduce la probabilidad de que el éxito o el fracaso de un estudiante dependa de su rendimiento en un

único tipo de instrumento evaluativo. Además, permite trayectorias de aprendizaje personalizadas, donde el monitoreo constante del progreso permite ofrecer apoyo específico y andamiaje a aquellos estudiantes que lo necesiten, en el momento en que lo necesiten, cerrando brechas antes de que se vuelvan insalvables.

No obstante, esta visión optimista debe ser matizada por una consideración crítica de los riesgos. La implementación ingenua de la evaluación por competencias puede, paradójicamente, generar nuevas formas de inequidad. Por ejemplo, si los criterios de una rúbrica están redactados con un lenguaje culturalmente sesgado o si valoran formas de comunicación más propias de un grupo dominante, pueden penalizar inadvertidamente a estudiantes de orígenes diversos.

De manera similar, las tareas de desempeño auténtico, como los proyectos de investigación, pueden requerir recursos (tiempo, acceso a materiales, capital social) que no están distribuidos de manera equitativa entre todo el estudiantado. Si el sistema no tiene en cuenta estas posibles disparidades, corre el riesgo de reemplazar una forma de sesgo por otra. La subjetividad en la evaluación, aunque mitigada por instrumentos como las rúbricas, también puede ser una fuente de sesgo implícito si los evaluadores no reciben una formación adecuada en equidad y conciencia cultural (Marion et al., 2020).

La mitigación de estos riesgos exige un compromiso proactivo y sistémico con el diseño de la equidad. Marion et al. (2020) proponen un marco de acción que se fundamenta en cuatro pilares. Primero, **involucrar a la comunidad** (incluyendo a estudiantes, familias y empleadores) en la definición de lo que constituye el éxito, para asegurar que los resultados de aprendizaje sean culturalmente relevantes y representativos. Segundo, **asegurar expectativas** consistentes a través de rúbricas claras y criterios de dominio transparentes, garantizando que todos los estudiantes comprendan qué se espera de ellos y cómo serán evaluados. Tercero, **monitorear y responder al progreso estudiantil**, utilizando los datos

de la evaluación formativa para identificar patrones de inequidad y para proporcionar apoyo diferenciado. Y cuarto, **adaptar continuamente el sistema**, utilizando procesos de mejora cíclica para refinar las prácticas evaluativas en función de la evidencia sobre su impacto en la equidad.

En conclusión, la evaluación por competencias no es una panacea para las inequidades en la educación superior, pero sí representa una de las herramientas más prometedoras para avanzar hacia una mayor justicia social en las aulas de ciencias exactas. La evidencia empírica demuestra su potencial para cerrar brechas de rendimiento para grupos subrepresentados al desacoplar el éxito académico de los predictores tradicionales y al ofrecer múltiples y más flexibles vías para demostrar el dominio.

Sin embargo, este potencial solo se realiza si la equidad se concibe no como un resultado deseable, sino como un principio de diseño fundamental que impregna cada aspecto del sistema evaluativo, desde la definición de las competencias hasta la formación de los evaluadores. La construcción de un sistema de este tipo no es solo una responsabilidad individual del docente, sino que requiere un compromiso a nivel de la cultura disciplinar y la política institucional, una dimensión que será el foco del siguiente y último apartado de este capítulo.

3.4. Impacto en la cultura disciplinar y la institución universitaria

Culminando el análisis sobre las consecuencias de la evaluación, la atención se desplaza ahora desde el nivel individual y grupal hacia la escala macro de la cultura disciplinar y la estructura institucional. La transición hacia un paradigma de evaluación por competencias no es un mero ajuste técnico que ocurre en el vacío del aula; es una intervención que choca, negocia y potencialmente transforma las normas, valores y prácticas profundamente arraigadas que constituyen la identidad de una disciplina y el funcionamiento de la universidad. El impacto a este nivel es complejo y a menudo conflictivo, ya que la evaluación por competencias desafía tradiciones largamente establecidas y exige una reconfiguración de las

políticas, los recursos y las relaciones de poder dentro de la institución. Investigaciones etnográficas y cualitativas han comenzado a desvelar la naturaleza de estas tensiones, identificando las culturas dominantes y las barreras sistémicas que condicionan la adopción de nuevos modelos evaluativos, particularmente en las ciencias exactas (Ebrahimpour, 2023; Sachmpazidi et al., 2024).

Esta confrontación cultural se manifiesta, en primera instancia, a través de una considerable **resistencia por parte del profesorado**. Esta resistencia no debe ser simplistamente atribuida a una mera aversión al cambio, sino que responde a una constelación de factores individuales, institucionales y culturales. Un estudio cualitativo de Ebrahimpour (2023) con 27 profesores universitarios identificó barreras individuales clave, como la falta de familiaridad con los nuevos modelos, la percepción de que son excesivamente complejos y la preocupación por un aumento de la carga de trabajo. De manera más profunda, emerge el temor a una pérdida de la autoridad tradicional en el aula, ya que el rol de transmisor de conocimiento se ve desplazado por el de facilitador. A nivel cultural, la adherencia a una identidad académica tradicional y la presión de los pares para mantener los estándares evaluativos convencionales actúan como poderosos inhibidores del cambio. La evaluación por competencias, al ser percibida como “subjetiva” o “no estructurada”, puede ser vista como una amenaza al rigor que define la identidad de la disciplina.

Precisamente, la **cultura disciplinar específica de la física y la matemática** presenta un conjunto particular de desafíos. El análisis etnográfico de Sachmpazidi et al. (2024) sobre los departamentos de física revela la existencia de una “cultura dominante” en torno a la evaluación, caracterizada por un fuerte énfasis en la resolución de problemas algorítmicos, un marcado individualismo y una cultura de competencia. Estas normas disciplinares implícitas, que definen qué cuenta como “conocimiento válido”, pueden entrar en conflicto directo con los principios de la evaluación por competencias. Por ejemplo, el énfasis en el aprendizaje colaborativo que a menudo acompaña a las metodologías de

proyectos choca con una tradición de evaluación individual y competitiva. De igual modo, la valoración de competencias holísticas y procesuales puede ser percibida por algunos como una “dilución” del rigor matemático y la precisión conceptual que se consideran el sello distintivo de estas disciplinas.

Para que el cambio sea posible, se requiere una profunda **transformación a nivel institucional**. La resistencia docente no puede ser superada únicamente con persuasión; necesita ser abordada mediante cambios estructurales y políticas universitarias coherentes. El estudio de Ebrahimpour (2023) destaca la importancia de las “restricciones institucionales y estructurales” como una barrera fundamental. Políticas poco claras sobre la evaluación, la falta de recursos para implementar nuevas metodologías, un desarrollo profesional inadecuado y, de manera crucial, criterios de promoción y titularidad que no valoran la innovación pedagógica, crean un entorno que desincentiva activamente el cambio. Si la institución recompensa exclusivamente la producción de investigación en detrimento de la excelencia y la innovación en la enseñanza y la evaluación, el profesorado no tendrá ningún incentivo real para invertir el tiempo y el esfuerzo que requiere la transición hacia un modelo competencial.

Esta necesaria reconfiguración institucional implica, por tanto, una serie de acciones coordinadas. En primer lugar, la universidad debe desarrollar y comunicar una **política de evaluación clara y coherente** que respalde explícitamente el enfoque por competencias. Esto proporciona al profesorado la seguridad y el marco necesarios para experimentar con nuevas prácticas. En segundo lugar, es indispensable una inversión significativa en **desarrollo profesional docente**. Como se mencionó en la sección 2.4, el rol de evaluador-mediador requiere un conjunto de competencias específicas que deben ser formadas y practicadas. Los talleres, las comunidades de práctica y los programas de mentoría son esenciales para construir la autoeficacia del profesorado y para superar la percepción de que los nuevos modelos son inmanejables (Ebrahimpour, 2023). Se analizaron los datos, se contrastaron las teorías, y la conclusión

es que el apoyo institucional es la variable clave que modula la capacidad de cambio del cuerpo docente.

Asimismo, la transición exige una reflexión sobre la **asignación de recursos y la gestión del tiempo**. La evaluación por competencias, especialmente aquella que implica retroalimentación detallada y el seguimiento de proyectos, es indudablemente más intensiva en tiempo que la corrección de un examen de opción múltiple. Las instituciones deben reconocer esta realidad y ajustar las cargas de trabajo docente, la ratio de estudiantes por profesor y el apoyo de personal asistente para que la implementación de una evaluación de alta calidad sea viable y sostenible. Ignorar esta dimensión material es condenar la reforma al fracaso, generando agotamiento y cinismo en el profesorado más comprometido (Sachmpazidi et al., 2024).

Finalmente, el impacto de la evaluación por competencias puede llegar a transformar la **identidad de la propia disciplina** a nivel departamental. Al cambiar lo que se evalúa, se cambia implícitamente la definición de lo que significa “ser un físico” o “ser un matemático”. El énfasis se desplaza desde la figura del solucionador solitario de problemas abstractos hacia la del profesional colaborativo, comunicador eficaz y pensador crítico capaz de aplicar su conocimiento a problemas complejos del mundo real. Este cambio puede enriquecer la disciplina, haciéndola más atractiva y accesible para una población estudiantil más diversa, como se sugirió en la sección 3.3. Sin embargo, también requiere que las comunidades disciplinares se abran a un diálogo franco sobre sus propias tradiciones y estén dispuestas a evolucionar. El estudio de Sachmpazidi et al. (2024) es un llamado a que los departamentos de física reconozcan y reflexionen críticamente sobre sus culturas evaluativas dominantes como un primer paso necesario para cualquier cambio significativo.

En conclusión, el impacto de la evaluación por competencias se extiende mucho más allá del aula, permeando la cultura de las disciplinas y la estructura de la universidad. La transición hacia este paradigma no es

un simple cambio de herramientas, sino una reforma sistémica que enfrenta resistencias culturales e institucionales significativas. Superar estas barreras requiere un enfoque multifacético que combine políticas institucionales claras, un desarrollo profesional robusto, una reasignación de recursos y, fundamentalmente, un diálogo abierto dentro de las propias comunidades disciplinares sobre su identidad y sus fines formativos. Al cerrar este capítulo sobre el impacto y las consecuencias, se completa el andamiaje teórico de esta obra, dejando el escenario preparado para el análisis del Capítulo 4, donde se examinará un estudio de caso empírico que ilustra cómo estas dinámicas se manifiestan en el contexto concreto de una universidad.

CAPÍTULO IV

EFECTOS DE LA EVALUACIÓN EDUCATIVA EN APRENDIZAJES POR COMPETENCIAS EN ESTUDIANTES DE MATEMÁTICA Y FÍSICA DE LA UNHEVAL DE HUÁNUCO

4.1. Introducción

Tras haber establecido en los capítulos precedentes el marco teórico, filosófico y metodológico que sustenta la transición hacia la evaluación por competencias, este capítulo se adentra en el terreno empírico. La validez de cualquier propuesta paradigmática reside, en última instancia, en su capacidad para ser observada, analizada y contrastada en contextos educativos reales. Por ello, esta sección presenta un estudio de caso que sirve como puente entre la teoría y la práctica: una investigación exhaustiva llevada a cabo en la Escuela Profesional de Matemática y Física de la Universidad Nacional Hermilio Valdizán (UNHEVAL) en Huánuco, Perú.

Este análisis localiza la discusión global sobre la crisis del modelo tradicional y la necesidad de un enfoque competencial, proporcionando una evidencia concreta de las dinámicas, desafíos y efectos de este cambio en una institución de educación superior específica. La literatura reciente

subraya la importancia de estos estudios situados, pues la evaluación por competencias no es un modelo monolítico, sino un proceso que debe adaptarse a las particularidades institucionales y disciplinares para favorecer la innovación y mejorar los perfiles de egreso (Menzala et al., 2023).

En este marco, la investigación en la UNHEVAL surge de una situación problemática claramente identificada: la coexistencia de un diseño curricular formalmente orientado por competencias con prácticas evaluativas que, en gran medida, permanecían ancladas en un enfoque tradicional. Este desfase, descrito por los propios docentes como una fuente de preocupación, situaba a la evaluación como el componente curricular más rezagado en el proceso de transformación pedagógica.

La persistencia de una evaluación centrada en la acumulación de contenidos fácticos y la reproducción memorística generaba una tensión insostenible con el objetivo de formar profesionales capaces de aplicar sus conocimientos de manera integral y contextualizada. Esta situación no es exclusiva de esta institución—se alinea con una tendencia documentada a nivel global (Menzala et al., 2023)—pero su manifestación local exigía una indagación específica para comprender sus efectos directos en los aprendizajes de los estudiantes de matemática y física (Cámara Acero et al., 2024).

Esta situación problemática se ve agravada por la complejidad inherente al contexto universitario, donde factores como el bienestar emocional de los estudiantes interactúan de formas no siempre predecibles con el rendimiento académico. Investigaciones en contextos similares han evidenciado que no existe necesariamente una correlación estadísticamente significativa entre el bienestar emocional y el desempeño académico, lo que sugiere que ambas dimensiones requieren estrategias de apoyo diferenciadas y una consideración cuidadosa dentro del modelo educativo (Sinchigalo et al., 2022). Este hallazgo subraya la necesidad de examinar cómo diferentes modelos evaluativos impactan no solo en la

adquisición de conocimientos, sino en una gama más amplia de factores formativos. La evaluación tradicional, con su énfasis en la calificación sumativa y los exámenes de alto riesgo, puede exacerbar la ansiedad y afectar negativamente el bienestar, mientras que un modelo basado en competencias, con su enfoque formativo y procesual, podría teóricamente contribuir a una experiencia de aprendizaje más positiva y holística (Menzala et al., 2023).

Frente a esta coyuntura, la realización de un estudio empírico en la UNHEVAL se justificó desde múltiples dimensiones. En el plano práctico, la investigación respondía a la necesidad imperiosa de la facultad de ciencias de la educación de diagnosticar y solucionar los problemas asociados a sus prácticas evaluativas para mejorar el desarrollo de competencias en sus estudiantes. En el ámbito teórico, el estudio se propuso sistematizar sus resultados para contribuir al campo gnoseológico de la ciencia educativa, demostrando la relación constante entre las variables de evaluación y aprendizaje por competencias.

Finalmente, su justificación metodológica residía en el desarrollo y validación de instrumentos de recolección de datos que, una vez demostrada su fiabilidad, podrían ser reutilizados en futuras investigaciones, fortaleciendo así la capacidad investigativa de la institución (Cámara Acero et al., 2024). Se analizaron los datos, se contrastaron las teorías, y la investigación se perfiló como una respuesta necesaria y multifacética a un problema local con resonancias globales.

Por consiguiente, para abordar esta problemática de manera sistemática, se llevó a cabo una investigación cuyos detalles metodológicos, hallazgos y discusión se presentarán en las secciones subsiguientes de este capítulo. El propósito de este análisis empírico fue explicar de manera rigurosa los efectos de la evaluación educativa en los aprendizajes por competencias de los estudiantes de matemática y física. A través de un diseño específico, se buscó determinar el grado de influencia de las prácticas evaluativas sobre el aprendizaje conceptual, procedimental y actitudinal, ofreciendo

así una radiografía detallada del estado de la cuestión en la UNHEVAL y generando evidencia para orientar futuras intervenciones pedagógicas.

Para guiar esta indagación y asegurar un abordaje estructurado del problema, el equipo de investigación estableció un conjunto claro de objetivos e hipótesis que enmarcaron todo el proceso. Estos elementos, que definieron el alcance y la dirección del estudio, se presentan a continuación de manera explícita, tal como fueron formulados en el proyecto original.

Objetivos de la Investigación

- **Objetivo General:** Explicar los efectos de la evaluación educativa en los aprendizajes por competencias en estudiantes de matemática y física de la UNHEVAL de Huánuco, periodo 2024.
- **Objetivos Específicos:**
 - a. Demostrar los efectos de la evaluación educativa en el aprendizaje conceptual en estudiantes del séptimo ciclo de matemática y física de la UNHEVAL de Huánuco, periodo 2024.
 - b. Demostrar los efectos de la evaluación educativa en el aprendizaje procedimental en estudiantes del séptimo ciclo de matemática y física de la UNHEVAL de Huánuco, periodo 2024.
 - c. Demostrar los efectos de la evaluación educativa en el aprendizaje actitudinal en estudiantes del séptimo ciclo de matemática y física de la UNHEVAL de Huánuco, periodo 2024.

Sistema de Hipótesis

- **Hipótesis General:** La evaluación educativa afecta significativamente los aprendizajes por competencias en estudiantes de matemática y física de la UNHEVAL de Huánuco, periodo 2024.

- **Hipótesis Específicas:**

H₁: La evaluación educativa afecta significativamente el aprendizaje conceptual en estudiantes del séptimo ciclo de matemática y física de la UNHEVAL de Huánuco, periodo 2024.

H₂: La evaluación educativa afecta significativamente el aprendizaje procedimental en estudiantes del séptimo ciclo de matemática y física de la UNHEVAL de Huánuco, periodo 2024.

H₃: La evaluación educativa afecta significativamente el aprendizaje actitudinal en estudiantes del séptimo ciclo de matemática y física de la UNHEVAL de Huánuco, periodo 2024.

4.2. Aspectos metodológicos

Para responder a la interrogante y los objetivos planteados en la sección anterior, la investigación desarrollada en la Universidad Nacional Hermilio Valdizán (UNHEVAL) se fundamentó en una arquitectura metodológica precisa y rigurosa, diseñada para garantizar la validez y fiabilidad de sus hallazgos. Este andamiaje empírico, que se describe en detalle en el presente apartado, se concibió para trascender la mera descripción de percepciones y adentrarse en la explicación de las relaciones entre las prácticas evaluativas y el desarrollo de competencias en los estudiantes.

La construcción de este tipo de estudios empíricos es fundamental en el campo, pues, como señalan investigaciones recientes, permite identificar

las brechas de aprendizaje y analizar datos de desempeño para promover mejoras continuas, lo que requiere un diseño metodológico robusto y sistemático (Garay et al., 2024). La elección de cada componente del diseño, desde el nivel de la investigación hasta los procedimientos de validación de instrumentos, se alineó con el propósito de generar evidencia empírica sólida sobre los efectos del modelo evaluativo en un contexto universitario real de ciencias exactas.

En consonancia con este propósito, el estudio se situó en un **nivel explicativo** de investigación. A diferencia de un estudio descriptivo, que se limitaría a caracterizar las variables, el enfoque explicativo se orientó a dilucidar las relaciones de causa-efecto entre la variable independiente (evaluación educativa) y la variable dependiente (aprendizaje por competencias).

Este nivel de profundidad busca responder no solo al “qué” sucede, sino al “porqué” de los fenómenos observados, explorando cómo las variaciones en las prácticas evaluativas influyen en los distintos dominios del aprendizaje estudiantil. En cuanto a su tipo, la investigación se clasificó como **básica o teórica**, ya que su intención principal no era la solución inmediata de un problema práctico, sino la generación de conocimiento científico que, posteriormente, pudiera ser aplicado para la mejora de las prácticas pedagógicas. Finalmente, se adoptó un **diseño causal explicativo**, una estructura metodológica que permite determinar el grado de influencia de una variable sobre otra en un contexto comparativo, lo cual resultó idóneo para contrastar las hipótesis formuladas sobre los efectos de la evaluación (Cámara Acero et al., 2024).

La concreción de este diseño se materializó en una población y una muestra cuidadosamente delimitadas. La **población** objeto de estudio estuvo constituida por la totalidad de los actores involucrados en el proceso formativo de la Escuela Profesional de Matemática y Física durante el semestre académico 2024-I, conformada por 60 estudiantes y 16 docentes. A partir de esta población, se seleccionó una **muestra** de

38 estudiantes del quinto año (noveno ciclo) y 13 docentes de la misma carrera profesional.

La selección de esta muestra se realizó mediante un **muestreo no probabilístico por conveniencia**, una técnica en la cual los sujetos no son elegidos al azar, sino de manera intencional por parte del investigador, atendiendo a su accesibilidad y a su pertinencia para los fines del estudio. En este caso, se seleccionó a los estudiantes del último año por considerar que habían estado expuestos a una gama más amplia y consolidada de prácticas evaluativas a lo largo de su formación, lo que los convertía en informantes clave para valorar los efectos acumulativos de dichas prácticas (Cámara Acero et al., 2024).

El núcleo del andamiaje investigativo se estructuró en torno a dos variables centrales: la **evaluación educativa** como variable independiente y el **aprendizaje por competencias** como variable dependiente. Para asegurar una medición precisa, ambos constructos latentes fueron sometidos a un riguroso proceso de **operacionalización**, descomponiéndolos en dimensiones e indicadores observables y medibles. La variable “evaluación educativa” se desglosó en cinco dimensiones: finalidad y función, extensión, agentes evaluadores, momento de aplicación y criterio de comparación.

Por su parte, la variable “aprendizaje por competencias” se operacionalizó en las tres dimensiones clásicas del saber: aprendizaje de contenidos conceptuales (saber conocer), aprendizaje de contenidos procedimentales (saber hacer) y aprendizaje de contenidos actitudinales (saber ser). Este proceso de descomposición es una práctica metodológica crucial en la evaluación de competencias, pues aborda el desafío de integrar múltiples dominios del aprendizaje (cognitivo, psicomotor, afectivo) en un marco de análisis coherente (Gallardo, 2020). Cada una de estas dimensiones e indicadores fue detallada en las matrices de operacionalización correspondientes (Cámara Acero et al., 2024).

Para capturar empíricamente estas variables operacionalizadas, se diseñaron y aplicaron **técnicas e instrumentos específicos**. La técnica principal fue la **encuesta**, y el instrumento utilizado consistió en dos **cuestionarios estructurados** con ítems de respuesta cerrada tipo **escala de Likert**. Se elaboró un cuestionario para cada grupo de la muestra: uno dirigido a los 13 docentes, diseñado para recabar datos sobre la variable independiente (sus prácticas y concepciones sobre la evaluación educativa), y otro dirigido a los 38 estudiantes, enfocado en medir la variable dependiente (su percepción sobre el desarrollo de sus aprendizajes por competencias). La elección de cuestionarios estructurados y escalas tipo Likert es coherente con diseños de investigación-acción y estudios cuantitativos que buscan medir la efectividad de herramientas y metodologías evaluativas en contextos de educación superior, ya que permiten una recolección de datos sistemática y su posterior análisis estadístico (Garay et al., 2024).

La rigurosidad de cualquier estudio cuantitativo depende críticamente de la calidad de sus instrumentos de medición. Por ello, la investigación de la UNHEVAL dedicó una atención especial a los **procedimientos de validez y confiabilidad**. En una primera fase, para asegurar la **validez de contenido** de los cuestionarios, estos fueron sometidos al **juicio de cinco expertos** con conocimiento en el área, quienes evaluaron la pertinencia, claridad y coherencia de cada ítem en relación con las variables y dimensiones que pretendían medir.

Se estableció como criterio que un instrumento sería considerado válido si obtenía una aprobación igual o superior al 80% por parte del panel de expertos. En una segunda fase, una vez validado el contenido, se procedió a determinar la **confiabilidad** de los instrumentos, es decir, su consistencia interna. Para este fin, se realizó una prueba piloto y se calculó el **coeficiente Alfa de Cronbach**, un estadístico que estima el

grado en que los ítems de un test se correlacionan entre sí. Se analizaron los datos, se contrastaron las teorías, y la aplicación de este coeficiente se alinea con las mejores prácticas internacionales, donde valores superiores a 0.80 son considerados un indicador de alta fiabilidad para instrumentos de evaluación educativa (Gallardo, 2020).

En síntesis, la arquitectura metodológica desplegada en el estudio de la UNHEVAL se caracterizó por un enfoque cuantitativo, un diseño causal-explicativo y el uso de instrumentos estructurados y validados que permitieron una recolección sistemática de datos sobre las variables de interés. La combinación de un muestreo intencional, una operacionalización detallada de las variables y un riguroso proceso de validación psicométrica conformó un andamiaje empírico sólido. Este conjunto sistemático de procedimientos investigativos, que materializa en la práctica los conceptos de diseño evaluativo discutidos en el Capítulo 2 de esta obra, proporciona la base de credibilidad necesaria para presentar e interpretar los hallazgos que se derivaron de su aplicación, los cuales serán objeto de análisis en la siguiente sección.

4.3. Hallazgos

La aplicación de los instrumentos metodológicos descritos en el apartado anterior permitió la recolección de un conjunto de datos empíricos que ofrecen una visión detallada sobre las dinámicas de la evaluación y el aprendizaje en la Escuela Profesional de Matemática y Física de la UNHEVAL. Esta sección se dedica a la presentación objetiva y organizada de dichos resultados, extraídos directamente del estudio de Cámara Acero et al. (2024). Los hallazgos se estructuran en dos bloques principales: en primer lugar, los resultados correspondientes a la variable “Evaluación Educativa”, que reflejan la perspectiva del personal docente; y en segundo lugar, los datos relativos a la variable “Aprendizaje por Competencias”, que recogen la percepción del estudiantado. La

presentación de estos datos brutos se realiza de manera descriptiva, sin incursionar en su interpretación o discusión, labor que se reserva para la sección subsecuente.

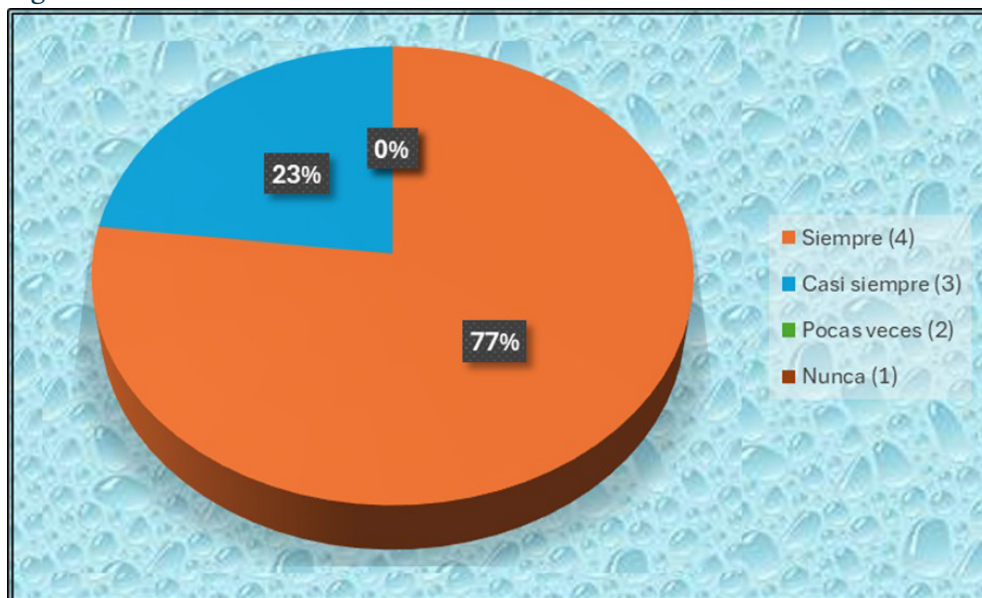
El análisis de los datos recabados del cuestionario aplicado a los 13 docentes de la muestra revela una percepción mayoritariamente favorable hacia la implementación de prácticas evaluativas alineadas con el enfoque por competencias. Los resultados generales, sintetizados en la Tabla 6, evidencian que una porción considerable del cuerpo docente manifiesta aplicar de manera constante este tipo de evaluación.

Tabla 6 Distribución de Frecuencias sobre la Percepción Docente de la Evaluación Educativa (General)

Escala Valorativa	Frecuencia (f)	Porcentaje (%)
Siempre (4)	09	69
Casi siempre (3)	04	31
Pocas veces (2)	00	00
Nunca (1)	00	00
Total	13	100

Esta tendencia general se puede visualizar de manera más clara en la Figura 1, que ilustra gráficamente la distribución porcentual de las respuestas docentes.

Figura 1 Niveles de la evaluación educativa.



Fuente: Elaboración propia

Desglosando estos resultados generales en sus componentes dimensionales, se observa una consistencia notable en la percepción docente a través de las distintas facetas de la evaluación. En la dimensión de **Finalidad y Función**, los datos indican una fuerte inclinación hacia el uso de la evaluación como un mecanismo de monitoreo y mejora del proceso de enseñanza-aprendizaje, tal como se detalla en la Tabla 7.

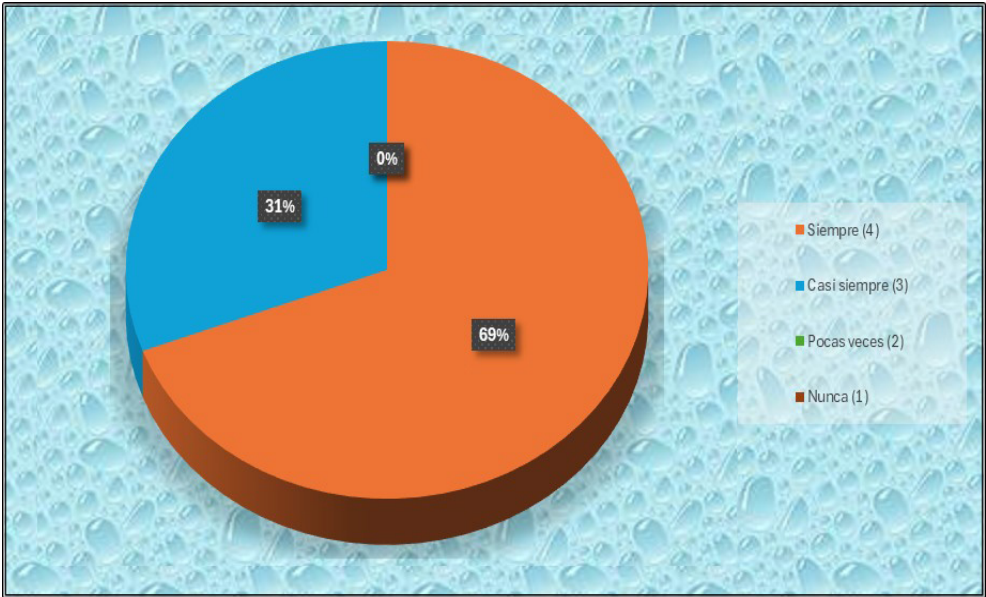
Tabla 7 Distribución de Frecuencias de la Dimensión Finalidad y Función

Escala Valorativa	Frecuencia (f)	Porcentaje (%)
Siempre (4)	10	77
Casi siempre (3)	03	23
Pocas veces (2)	00	00
Nunca (1)	00	00
Total	13	100

Fuente: Tabla elaborada a partir de los datos de la “Tabla 3” de Cámara Acero et al. (2024, p. 47).

La representación gráfica de estos porcentajes, que debería incluirse como Figura 2, confirma visualmente esta tendencia.

Figura 2 Niveles de la dimensión finalidad y función



Fuente: Elaboración propia

Un patrón idéntico se replica en las demás dimensiones analizadas. Por ejemplo, en lo que respecta a la dimensión Momento de Aplicación, los hallazgos demuestran una práctica declarada de evaluación continua a lo largo de todo el proceso formativo, abarcando evaluaciones de entrada, de proceso y de salida, como se muestra en la Tabla 8.

Tabla 8 Distribución de Frecuencias de la Dimensión Momento de Aplicación

Escala Valorativa	Frecuencia (f)	Porcentaje (%)
Siempre (4)	10	77
Casi siempre (3)	03	23
Pocas veces (2)	00	00
Nunca (1)	00	00
Total	13	100

Fuente: Tabla elaborada a partir de los datos de la “Tabla 6” de Cámara Acero et al. (2024, p. 49).

Al examinar los datos desde la perspectiva de los 38 estudiantes encuestados, se obtiene una panorámica de su percepción sobre el logro de aprendizajes en el marco del enfoque por competencias. Los resultados generales, presentados en la Tabla 9, indican que la mayoría del estudiantado percibe que logra asimilar los contenidos y desarrollar las competencias propuestas en sus asignaturas.

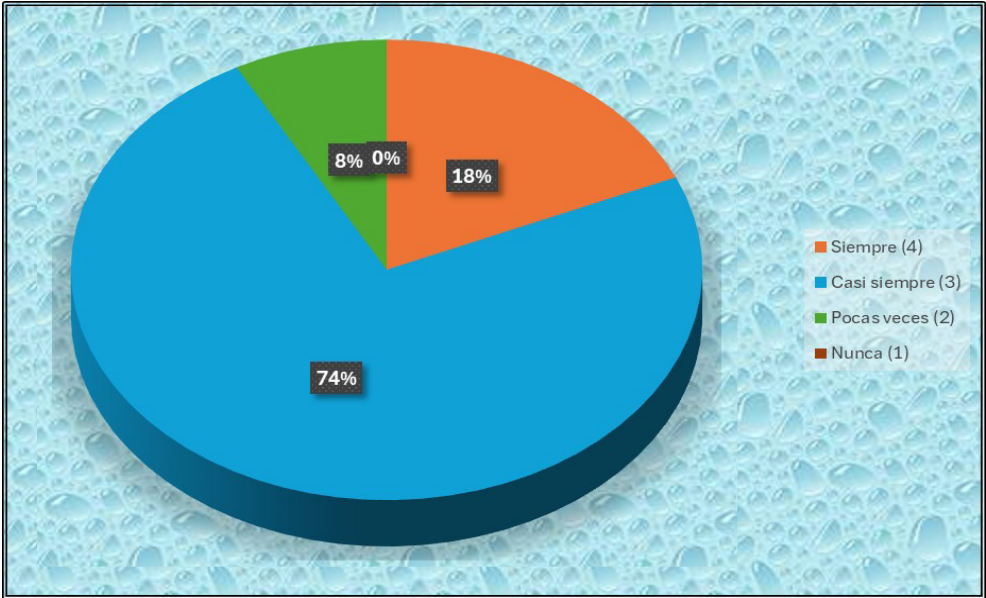
Tabla 9 Distribución de Frecuencias sobre el Aprendizaje por Competencias (General)

Escala Valorativa	Frecuencia (f)	Porcentaje (%)
Siempre (4)	07	18
Casi siempre (3)	28	74
Pocas veces (2)	03	08
Nunca (1)	00	00
Total	38	100

Fuente: Tabla elaborada a partir de los datos de la “Tabla 9” de Cámara Acero et al. (2024, p. 52).

La Figura 3 ofrece una visualización de esta distribución, destacando el predominio de las valoraciones positivas por parte de los estudiantes.

Figura 3 Niveles de frecuencias de los aprendizajes por competencias



Fuente: Elaboración propia

El análisis de estos datos por dimensiones revela matices en la percepción estudiantil. En la dimensión de **Contenido Conceptual**, que evalúa la asimilación de teorías y significados, los resultados se distribuyen como se muestra en la Tabla 10.

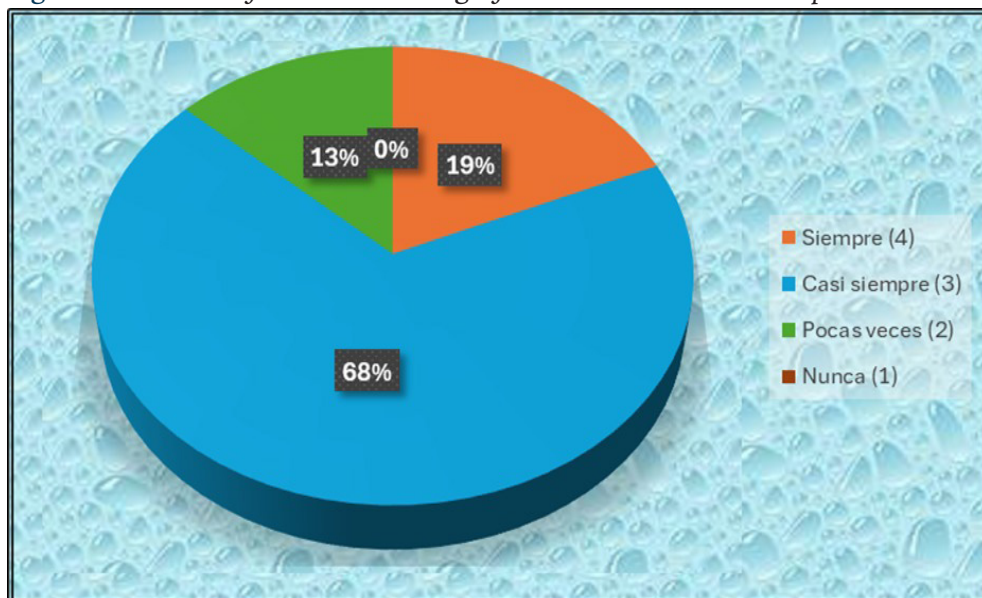
Tabla 10 Distribución de Frecuencias de la Dimensión Contenido Conceptual

Escala Valorativa	Frecuencia (f)	Porcentaje (%)
Siempre (4)	07	19
Casi siempre (3)	26	68
Pocas veces (2)	05	13
Nunca (1)	00	00
Total	38	100

Fuente: Tabla elaborada a partir de los datos de la “Tabla 10” de Cámara Acero et al. (2024, p. 52).

La correspondiente Figura 4 permitiría observar esta distribución porcentual de manera gráfica.

Figura 4 Niveles de frecuencias del significado del contenido conceptual



Fuente: Elaboración propia

En cuanto a la dimensión de **Contenido Procedimental**, que se refiere al desarrollo de habilidades y técnicas, los resultados muestran una distribución ligeramente diferente, la cual se detalla en la “Tabla 11” de Cámara Acero et al. (2024, p. 53) y se visualiza en la “Figura 12” de la misma página. Finalmente, la dimensión de **Contenido Actitudinal**, que aborda la internalización de valores y una base ética, presenta los resultados que se condensan en la Tabla 11.

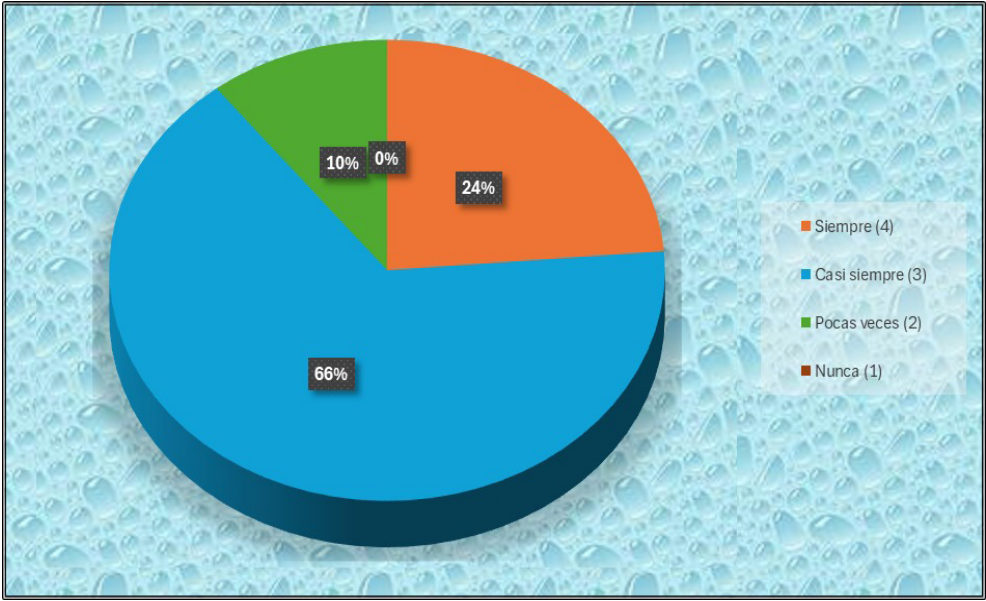
Tabla 11 Distribución de Frecuencias de la Dimensión Contenido Actitudinal

Escala Valorativa	Frecuencia (f)	Porcentaje (%)
Siempre (4)	09	24
Casi siempre (3)	25	66
Pocas veces (2)	04	10
Nunca (1)	00	00
Total	38	100

Fuente: Tabla elaborada a partir de los datos de la “Tabla 13” de Cámara Acero et al. (2024, p. 54).

La representación visual de estos últimos datos se encuentra en la Figura 5

Figura 5 Niveles de frecuencias del significado del contenido actitudinal



Fuente: Elaboración propia

Estos hallazgos cuantitativos, tanto desde la perspectiva docente como estudiantil, conforman la base empírica sobre la cual se construirán la discusión y las conclusiones de este estudio de caso en la sección que sigue.

4.4. Discusión y conclusiones

El conjunto de datos empíricos presentados en la sección anterior ofrece una panorámica detallada de las dinámicas evaluativas en la Universidad Nacional Hermilio Valdizán (UNHEVAL), proporcionando una base sólida para la interpretación y el contraste con el marco teórico de la presente obra. A partir del análisis inferencial de estos hallazgos, el estudio de Cámara Acero et al. (2024) arribó a una serie de conclusiones que no solo responden a las hipótesis planteadas, sino que también iluminan la compleja relación entre las prácticas de evaluación y el desarrollo de competencias en el contexto específico de la formación de docentes de matemática y física. La discusión de estos resultados, contextualizada tanto en la literatura académica pertinente como en el andamiaje conceptual de este libro, permite trascender el caso particular de la UNHEVAL para extraer implicaciones de mayor alcance para la educación superior en ciencias. Este proceso de diálogo entre la evidencia empírica y la teoría es fundamental, pues, como demuestran estudios sobre la implementación de la evaluación por competencias, la objetividad del proceso se enriquece al triangular múltiples perspectivas y datos, aunque ello incrementa su complejidad (Sanz et al., 2023).

La conclusión central y más abarcadora de la investigación es que **la evaluación educativa influye de manera significativa en el aprendizaje por competencias de los estudiantes** del quinto año de la Escuela Profesional de Matemática y Física. Este hallazgo general, que confirma la hipótesis principal del estudio, se desagrega en tres conclusiones específicas que corroboran, a su vez, las hipótesis secundarias. Primero, se constató que la evaluación educativa influye en el **aprendizaje conceptual de los estudiantes**. Segundo, se evidenció una influencia significativa de la **evaluación educativa en el aprendizaje procedimental**. Y, tercero, se demostró que la **evaluación educativa influye en el aprendizaje actitudinal de los futuros profesionales** (Cámara Acero et al., 2024). Estas cuatro conclusiones, tomadas en conjunto, proporcionan una validación empírica robusta a la premisa de que la forma en que se evalúa

no es un acto neutro o meramente administrativo, sino una intervención pedagógica con efectos directos y profundos en todas las dimensiones del saber.

Esta correlación positiva encuentra eco en la discusión que los propios autores de la investigación establecen con la literatura precedente. Los hallazgos de la UNHEVAL son consistentes con los de Olmos (2008), quien ya señalaba la utilidad de los procesos de evaluación formativa y el uso de la tecnología para mejorar la eficacia del aprendizaje en contextos universitarios. Asimismo, la capacidad de la evaluación para promover el aprendizaje por competencias, evidenciada en el caso peruano, se alinea con las conclusiones de Triadó et al. (2012), quienes destacaron la importancia de adaptar las metodologías docentes para facilitar la adquisición de competencias en el pregrado. En lo que respecta a las dimensiones específicas del aprendizaje, la influencia positiva sobre el componente actitudinal, documentada en la UNHEVAL, resuena con los hallazgos de Ferrándiz-Vindel (2011), quien subrayó el poder de la autoevaluación para mejorar la percepción del estudiante sobre su propio proceso de aprendizaje. Finalmente, la mejora en las prácticas evaluativas, implícita en los resultados positivos del estudio, establece un paralelo con la investigación de Lorenzana (2012), que exploró cómo el conocimiento y manejo de un sistema de evaluación por competencias incide directamente en la mejora de las prácticas docentes (Cámara Acero et al., 2024).

Más allá del diálogo con la literatura existente, la significancia del estudio de la UNHEVAL se magnifica cuando se interpreta a la luz del marco teórico-filosófico desarrollado en este libro. La investigación empírica en Huánuco sirve como una **instanciación concreta de la crisis del modelo tradicional** y la necesidad de una ruptura paradigmática, tal como se argumentó en el Capítulo 1. El problema que motivó el estudio — la brecha entre un currículo por competencias y unas prácticas evaluativas ancladas en la tradición— es la manifestación en el mundo real de la tensión epistemológica discutida en la sección 1.3. Las conclusiones positivas del estudio, que demuestran que una evaluación (presumiblemente

más auténtica) influye favorablemente en el aprendizaje conceptual, procedimental y actitudinal, ofrecen un respaldo empírico a la idea de que **contenido y competencia no son opuestos, sino simbiotes**. Lejos de abandonar el rigor del contenido, el enfoque competencial parece potenciar su asimilación y aplicación en todas sus facetas.

La robustez de estas conclusiones depende, a su vez, de la solidez de su andamiaje metodológico, el cual también ilustra y valida los principios expuestos en el Capítulo 2 de esta obra. La decisión de Cámara Acero et al. (2024) de utilizar **instrumentos estructurados y someterlos a un riguroso proceso de validación por expertos y análisis de fiabilidad** (mediante el coeficiente Alfa de Cronbach) ejemplifica la necesidad de herramientas de valoración precisas, tal como se discutió en la sección 2.3. Se analizaron los datos, se contrastaron las teorías, y la construcción de cuestionarios específicos para medir tanto las prácticas evaluativas como los aprendizajes competenciales materializa en la práctica la necesidad de ir más allá de la subjetividad, utilizando instrumentos que permitan una recolección de datos sistemática. La investigación-acción llevada a cabo por Garay et al. (2024) en el campo de la ingeniería refuerza esta idea, mostrando que el desarrollo de herramientas de evaluación estructuradas es un paso indispensable para identificar brechas de aprendizaje y promover mejoras continuas.

Esta dimensión metodológica se complementa con una reflexión sobre el rol docente. Aunque el estudio de la UNHEVAL no midió directamente las competencias de los profesores, sus conclusiones positivas permiten inferir la existencia de un núcleo de docentes que ya actúan como **evaluadores-mediadores**, tal como se definió en la sección 2.4. La capacidad de un sistema evaluativo para influir positivamente en el aprendizaje actitudinal, por ejemplo, sugiere prácticas que van más allá de la simple calificación, implicando retroalimentación formativa y una preocupación por el desarrollo integral del estudiante. El hecho de que la evaluación formativa mediada por el docente pueda explicar hasta el 70.8% de la variabilidad en el desarrollo de competencias científicas

(Bellido et al., 2024) refuerza la idea de que los resultados positivos en la UNHEVAL son, en gran medida, un testimonio de la acción mediadora de su profesorado. La recomendación de la propia tesis de impulsar programas de capacitación docente (Cámara Acero et al., 2024, p. 58) es un reconocimiento implícito de que la efectividad de cualquier modelo evaluativo depende crucialmente de las competencias de quienes lo implementan.

Finalmente, la significancia del estudio de la UNHEVAL trasciende lo técnico y lo metodológico para conectar directamente con las **dimensiones humanas y sociales de la evaluación**, el tema central del Capítulo 3. La conclusión de que la evaluación influye significativamente en el “aprendizaje actitudinal” es, quizás, la más profunda. Proporciona una evidencia empírica directa para el argumento expuesto en la sección 3.2 sobre el **impacto de la evaluación en la identidad, la motivación y el bienestar del aprendiz**. Una práctica evaluativa que es percibida como justa, relevante y formativa contribuye a construir una identidad profesional sólida, una motivación intrínseca por el aprendizaje y una base ética para la futura práctica profesional. Esto se contrapone a las consecuencias humanas negativas de un modelo tradicional punitivo, que a menudo genera ansiedad y desvinculación (como se discutió en la sección 3.1). Los hallazgos del estudio de Sanz et al. (2023), que demuestran la complejidad pero también la mayor objetividad de los procesos de evaluación que involucran a múltiples agentes, apuntan en la misma dirección: un proceso evaluativo más participativo y dialógico, aunque más exigente, es percibido como más legítimo y formativo.

En síntesis, el estudio de caso de la Universidad Nacional Hermilio Valdizán cierra el ciclo argumentativo de esta obra, funcionando como un microcosmos donde las discusiones teóricas, filosóficas, metodológicas y humanas convergen y se validan. La investigación no solo responde afirmativamente a sus propias hipótesis sobre la influencia de la evaluación en el aprendizaje por competencias, sino que, al hacerlo, proporciona una evidencia empírica robusta para los argumentos centrales de este

libro. Demuestra que la crisis del modelo tradicional es una realidad palpable en las instituciones de educación superior; que la evaluación por competencias, lejos de ser un ideal abstracto, tiene efectos medibles y positivos en el aprendizaje conceptual, procedimental y actitudinal; que su implementación requiere de un diseño metodológico riguroso y de instrumentos validados; y que su éxito depende de manera fundamental de la figura del docente como un mediador experto. La experiencia de la UNHEVAL, por tanto, no es solo un reporte de hallazgos locales; es un testimonio elocuente de la necesidad y la viabilidad de la transformación paradigmática que este libro defiende.

REFLEXIONES FINALES

Al llegar al término de este recorrido analítico, se hace imperativo tejer los hilos argumentativos que se han desplegado a lo largo de la obra para construir una visión de conjunto, una síntesis que no solo recapitule, sino que trascienda las conclusiones parciales. Este libro partió de la constatación de una profunda crisis en el paradigma evaluativo de la educación superior científica. A través de un examen crítico, hemos argumentado que la solución a esta crisis no reside en meros ajustes instrumentales, sino en una re-arquitectura sistémica de la evaluación, fundamentada en el constructo de competencia. Esta tesis, planteada en la introducción, ha sido validada y enriquecida dialécticamente a través de los capítulos. Se ha demostrado que la insuficiencia epistemológica del modelo tradicional (Capítulo 1) exige la adopción de principios de evaluación auténtica (Capítulo 2), los cuales se materializan en metodologías como el Aprendizaje Basado en Proyectos y la evaluación por portafolios. Estas estrategias, a su vez, requieren de instrumentos de valoración rigurosos como las rúbricas analíticas y, de manera crucial, de la actuación de un docente que asuma un rol de evaluador-mediador. La investigación ha evidenciado que la efectividad de este andamiaje no es meramente teórica; tiene consecuencias tangibles en la dimensión humana y social del aprendizaje (Capítulo 3) y encuentra su correlato empírico en estudios de caso como el de la Universidad Nacional Hermilio Valdizán (Capítulo 4), que constata la influencia significativa de la evaluación en el desarrollo de competencias conceptuales, procedimentales y actitudinales (Cámara Acero et al., 2024).

Esta cadena argumental revela que la evaluación por competencias no es un fin en sí misma, sino el medio para alcanzar un objetivo superior: el cultivo del pensamiento científico. Al desplazar el foco de la medición de un saber inerte hacia la valoración de un saber en acción, se promueve una forma de aprendizaje más profunda, significativa y transferible. La simbiosis entre contenido y competencia, defendida en el Capítulo 1, se ha mostrado no como una aspiración idealista, sino como una realidad funcional. La competencia, lejos de diluir el contenido, lo resignifica y lo dota de propósito, tal como lo ilustra la distinción entre la sintaxis matemática y su significado físico (Redish & Kuo, 2015). El análisis empírico sobre las creencias epistemológicas refuerza esta idea, mostrando que la competencia matemática no depende de una visión particular sobre la certeza del conocimiento, sino de la capacidad argumentativa para justificar cualquier posición (Rott & Leuders, 2017). Se analizaron los datos, se contrastaron las teorías, y la conclusión es inequívoca: la evaluación por competencias, lejos de ser una amenaza al rigor, es su manifestación más auténtica.

La viabilidad de esta manifestación, sin embargo, depende de una arquitectura metodológica sólida. Los principios de contextualización, complejidad y carácter procesual (Konyeme & Alordiah, 2024) no son meras directrices abstractas, sino los cimientos sobre los que se deben edificar las estrategias de evaluación. Metodologías como el Aprendizaje Basado en Proyectos han demostrado empíricamente su capacidad para mejorar de manera sustancial las habilidades de resolución de problemas y el pensamiento crítico en física (Nicholus et al., 2023), mientras que la evaluación por portafolios se ha revelado como una herramienta potente para documentar el desarrollo de habilidades de proceso científico y fomentar la práctica reflexiva (Shidiq et al., 2024). No obstante, la efectividad de estas estrategias macro depende de la calidad de los instrumentos micro. La investigación ha evidenciado que es posible diseñar rúbricas analíticas con altos coeficientes de fiabilidad ($g = 0.85$) para valorar procesos cognitivos complejos como el pensamiento crítico o el razonamiento científico, garantizando la objetividad del proceso

(Reynders et al., 2020; Timmerman et al., 2011).

Esta transición sistémica, sin embargo, no puede ser reducida a una cuestión meramente técnica; es, ante todo, un desafío humano e institucional. La figura del docente como evaluador-mediador se erige como el pilar fundamental de todo el edificio. La evidencia cuantitativa es contundente: la acción mediadora del profesorado, manifestada en prácticas de evaluación formativa como la retroalimentación oportuna y la facilitación de la autoevaluación, puede llegar a explicar más del 70% de la variabilidad en el desarrollo de competencias científicas (Bellido et al., 2024). Este dato transforma la percepción del rol docente, de un transmisor de información y calificador a un arquitecto de experiencias de aprendizaje y un catalizador del desarrollo cognitivo y metacognitivo. Las competencias necesarias para este nuevo rol —diseño, facilitación, liderazgo y mediación socio-cognitiva— pueden y deben ser desarrolladas a través de una formación docente específica y continua (Guasch et al., 2010).

A partir de esta síntesis integrada, es posible elevar la mirada y explorar las implicaciones transversales que esta transformación evaluativa conlleva para la educación superior. En primer lugar, a nivel de **política institucional**, la adopción de la evaluación por competencias no puede ser una iniciativa aislada de docentes innovadores. Requiere un compromiso institucional explícito que se traduzca en cambios estructurales. Esto incluye la revisión de los sistemas de promoción y reconocimiento docente para valorar el tiempo y el esfuerzo que demanda el diseño y la implementación de la evaluación auténtica. Implica, también, una reasignación de recursos para apoyar la creación de materiales, el acceso a tecnologías de simulación y la reducción de la ratio estudiante-profesor en asignaturas que implementen estas metodologías. Asimismo, los sistemas de aseguramiento de la calidad deben evolucionar para medir no solo indicadores de resultado (tasas de egreso, calificaciones), sino también indicadores de proceso que capturen la calidad de las experiencias de aprendizaje y evaluación.

En segundo lugar, las implicaciones para la **formación del profesorado universitario** son profundas. La formación inicial y continua de los docentes de ciencias no puede seguir centrándose exclusivamente en la profundización del contenido disciplinar. Es imperativo integrar, de manera transversal y sistemática, una sólida formación pedagógica y didáctica centrada en la evaluación. Los programas de desarrollo profesional deben equipar a los docentes con las competencias para diseñar tareas auténticas, construir y validar rúbricas, proporcionar retroalimentación efectiva y gestionar dinámicas de co-evaluación. Esto no es un añadido, sino una competencia profesional esencial para el docente universitario del siglo XXI. La transformación evaluativa no ocurrirá por decreto; ocurrirá a través del empoderamiento y la capacitación del cuerpo docente.

En tercer lugar, este cambio tiene el potencial de transformar la **cultura disciplinar** de las ciencias exactas. Tradicionalmente, estas disciplinas han sido percibidas como dominios de conocimiento abstracto, jerárquico y a menudo intimidante. La evaluación por competencias, al centrarse en la resolución de problemas auténticos y en el trabajo colaborativo, puede contribuir a proyectar una imagen más humana, creativa y socialmente relevante de la física y la matemática. Fomenta una cultura de la indagación y la argumentación, en lugar de una cultura de la respuesta correcta. Al valorar el proceso y la reflexión, se promueve una visión de la ciencia no como un cuerpo de dogmas, sino como una forma de pensar y de construir conocimiento, lo cual puede, a su vez, atraer a un espectro más diverso de estudiantes a estas disciplinas.

A pesar de la robustez de los argumentos y las evidencias presentadas, es fundamental reconocer las **limitaciones de esta obra y los horizontes que deja abiertos**. Nuestro análisis se ha centrado primordialmente en las ciencias exactas, y aunque muchos de los principios y metodologías son potencialmente transferibles, su aplicación en las ciencias sociales, las humanidades o las artes requeriría de una contextualización y adaptación cuidadosa. La investigación futura debería explorar estas transferencias interdisciplinarias. Además, el estudio de caso presentado,

si bien es profundo, corresponde a una realidad institucional y cultural específica; se necesitan más estudios comparativos e internacionales para identificar patrones universales y variaciones contextuales. Asimismo, la mayoría de los estudios citados son de naturaleza transversal. Se requiere una agenda de investigación que impulse estudios longitudinales para comprender el impacto a largo plazo de la evaluación por competencias en el desarrollo profesional y la trayectoria vital de los egresados. Finalmente, la irrupción de la inteligencia artificial generativa plantea un nuevo y fascinante horizonte de investigación, con preguntas urgentes sobre cómo estas tecnologías pueden ser aprovechadas para potenciar la evaluación auténtica (por ejemplo, creando simulaciones más complejas o proveyendo retroalimentación personalizada a gran escala) y, al mismo tiempo, cuáles son los riesgos éticos y pedagógicos que debemos mitigar.

En última instancia, este libro ha buscado defender una idea fundamental: la evaluación no es solo lo que hacemos para medir el aprendizaje, sino que es, en sí misma, uno de los actos de aprendizaje más potentes que podemos diseñar. Repensar la evaluación en la formación científica universitaria es mucho más que una actualización técnica; es una cuestión de coherencia epistemológica, de eficacia pedagógica y, sobre todo, de responsabilidad ética. Se trata de asegurar que estamos formando no solo a individuos que conocen la ciencia, sino a profesionales y ciudadanos que pueden pensar y actuar científicamente, capaces de enfrentar con rigor, creatividad y sentido crítico los complejos desafíos de nuestro tiempo. La transición hacia una epistemología de la praxis evaluativa, centrada en la valoración de la competencia, es el camino ineludible para una educación superior que aspire a ser verdaderamente transformadora.

REFERENCIAS BIBLIOGRÁFICAS

- Bellido, R., Venturo, C., Cruzata, A., Sarmiento, E., Corro, J., & Rejas, L. (2024). Involvement of the student in their learning: Effects of formative assessment on competency development. *EURASIA Journal of Mathematics, Science and Technology Education*, 20(5), em2440. <https://www.ejmste.com/article/involvement-of-the-student-in-their-learning-effects-of-formative-assessment-on-competency-14453>
- Cámara, A, Rivera, E., Rojas, A., Trujillo, P., & Ascayo, O. (2024). *Efectos de la evaluación educativa en aprendizajes por competencias en estudiantes de matemática y física de la UNHEVAL de Huánuco* [Tesis de investigación]. Universidad Nacional Hermilio Valdizán.
- Castrillo, C., & Triminio, C. (2024). La evaluación por competencia en el contexto universitario de las carreras de Matemáticas y Física-Matemática. *Revista Científica Retos de la Ciencia*, 8(18), 55-72. <https://www.retosdelacienciaec.com/Revistas/index.php/retos/article/view/489>

- Chiu, M., & Lin, J. (2019). Modeling competence in science education. *Disciplinary and Interdisciplinary Science Education Research*, 1(1), 12. <https://link.springer.com/article/10.1186/s43031-019-0012-y>
- Ebrahimpour, D. (2023). Factors Contributing to Resistance Toward Alternative Assessment Models Among Faculty. *Assessment and Practice in Educational Sciences*, 1(3), 10-18. <https://journalapes.com/index.php/apes/article/view/20>
- Gallardo, K. (2020). Competency-based assessment and the use of performance-based evaluation rubrics in higher education: Challenges towards the next decade. *Problems of Education in the 21st Century*, 78(1), 61-79. <https://www.ceeol.com/search/article-detail?id=940597>
- Garay, C., Castillo, A., Gijón, C., Domínguez, G., Rosales, C., & Oliart, A. (2024). Competency-based assessment tools for engineering higher education: a case study on complex problem-solving. *Cogent Education*, 11(1), 2392424. <https://www.tandfonline.com/doi/abs/10.1080/2331186X.2024.2392424>
- Guasch, T., Alvarez, I., & Espasa, A. (2010). University teacher competencies in a virtual teaching/learning environment: Analysis of a teacher training experience. *Teaching and Teacher Education*, 26(2), 199-206. <https://www.sciencedirect.com/science/article/pii/S0742051X09000535>
- Heyder, A., & Pegels, H. (2025). Detrimental effects of instructors' fixed mindsets on students' anticipated motivation and emotions in secondary and higher education. *Social Psychology of Education*, 28(1), 32. <https://link.springer.com/article/10.1007/s11218-024-10001-9>

- Jones, E., Priestley, M., Brewster, L., Wilbraham, E., Hughes, G., & Spanner, L. (2021). Student wellbeing and assessment in higher education: The balancing act. *Assessment & Evaluation in Higher Education*, 46(3), 438-450. <https://www.tandfonline.com/doi/abs/10.1080/02602938.2020.1782344>
- KONYEME, J., & ALORDIAH, C. (2024). Rethinking assessment: The need for authentic evaluation in Science Education. *University of Delta Journal of Multidisciplinary Research*, 1(2), 68-81. <https://umrj.org.ng/wp-content/uploads/2024/09/RETHINKING-ASSESSMENT-THE-NEED-FOR-AUTHENTIC-EVALUATION-IN-SCIENCE-EDUCATION.pdf>
- Marion, S., Worthen, M., & Evans, C. (2020). *How systems of assessments aligned with competency-based education can support equity*. Aurora Institute. <https://eric.ed.gov/?id=ED603989>
- Menzala, C., Ortega, E., Menzala, R., & Solís, B. (2023). Evaluación basada en competencias en educación superior. *Horizontes Revista de Investigación en Ciencias de la Educación*, 7(28), 836-851. http://www.scielo.org.bo/scielo.php?pid=S2616-79642023000200836&script=sci_arttext
- Nicholus, G., Muwonge, C., & Joseph, N. (2023). The role of problem-based learning approach in teaching and learning physics: a systematic literature review. *F1000Research*, 12, 951. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10646338/>
- Nkopodi, N., Jakovljevic, M., & Photo, P. (2024). Criteria for enhancing student wellbeing in STEM classrooms: ICT and indigenous knowledge in South African higher education. *Discover Education*, 3(1), 152. <https://link.springer.com/article/10.1007/s44217-024-00251-2>

- Pérez-Jorge, D., Boutaba-Alehyar, M., González-Contreras, A. I., & Pérez-Pérez, I. (2025). Examining the effects of academic stress on student well-being in higher education. *Humanities and Social Sciences Communications*, 12(1), 1-13. <https://www.nature.com/articles/s41599-025-04698-y>
- Redish, E., & Kuo, E. (2015). Language of physics, language of math: Disciplinary culture and dynamic epistemology. *Science & Education*, 24(5), 561-590. <https://link.springer.com/article/10.1007/s11191-015-9749-7>
- Reynders, G., Lantz, J., Ruder, S., Stanford, C., & Cole, R. (2020). Rubrics to assess critical thinking and information processing in undergraduate STEM courses. *International Journal of STEM Education*, 7(1), 9. <https://link.springer.com/article/10.1186/s40594-020-00208-5>
- Richard, M., Delgado, J., LeGresley, S., & Fischer, C. (2022). *Implementing competency-based grading improves the performance of women and first generation students in introductory physics* [Preprint]. arXiv. <https://arxiv.org/abs/2206.10574>
- Rott, B., & Leuders, T. (2017). Mathematical competencies in higher education: Epistemological beliefs and critical thinking in different strands of pre-service teacher education. *Journal for Educational Research Online*, 9(2), 115-136. https://www.pedocs.de/frontdoor.php?source_opus=14899
- Sachmpazidi, D., Turpen, C., Petrella, J., Dalka, R., & Abdurrahman, F. (2024). Recognizing dominant cultures around assessment and educational change in physics programs. *Physical Review Physics Education Research*, 20(1), 010132. <https://journals.aps.org/prper/abstract/10.1103/PhysRevPhysEducRes.20.010132>

- Sanz, I., Lázaro, J., & Usart, M. (2023). Formar y evaluar competencias en educación superior: Una experiencia sobre inclusión digital. *RIED. Revista Iberoamericana de Educación a Distancia*, 26(2), 205-225. <https://www.re-dalyc.org/journal/3314/331474781010/331474781010.pdf>
- Shidiq, G., Bunnaen, W., & Widiyan, A. (2024). USING PORTFOLIO ASSESSMENT TO ENHANCE STUDENTS'SCIENCE PROCESS SKILLS IN STEM LEARNING ACTIVITIES. *Jurnal Pendidikan Matematika dan IPA*, 15(1), 105-118. <https://jurnal.untan.ac.id/index.php/PMP/article/view/72228/0>
- Sinchigalo, R., Guzmán, B., & Bonilla, D. (2022). Bienestar emocional y rendimiento académico en estudiantes universitarios: relación bidimensional y su impacto en las estrategias de apoyo. *Journal of science and research*, 7(4), 71-95. <https://revistas.utb.edu.ec/index.php/sr/article/view/2898>
- Thi, H., Thi, V., & Nguyen, T. (2024). Metacognition in mathematics education: From academic chronicle to future research scenario—A bibliometric analysis with the Scopus database. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(4), em2427. <https://www.ejmste.com/article/metacognition-in-mathematics-education-from-academic-chronicle-to-future-research-scenario-a-14381>
- Timmerman, B., Strickland, D., Johnson, R., & Payne, J. (2011). Development of a 'universal' rubric for assessing undergraduates' scientific reasoning skills using scientific writing. *Assessment & Evaluation in Higher Education*, 36(5), 509-547. <https://www.tandfonline.com/doi/abs/10.1080/02602930903540991>

- Tiruneh, D., De Cock, M., Weldelessie, A., Elen, J., & Janssen, R. (2017). Measuring critical thinking in physics: Development and validation of a critical thinking test in electricity and magnetism. *International Journal of Science and Mathematics Education*, 15(4), 663-682. <https://link.springer.com/article/10.1007/s10763-016-9723-0>
- Tüysüz, M., Tüzün, Ü., & Çavus, H. (2023). Designing Authentic Science Assessment for Pre-Service Science Teachers. *Dinamika Ilmu*, 23(2), 261-276. <https://eric.ed.gov/?id=EJ1410223>
- Vo, D., & Simmie, G. (2024). Assessing scientific inquiry: A systematic literature review of tasks, tools and techniques. *International Journal of Science and Mathematics Education*, 23(4), 871-906. <https://link.springer.com/article/10.1007/s10763-024-10498-8>

