



ISSN 1992-6510
e-ISSN 2520-9299



REALIDAD Y REFLEXIÓN ES UNA PUBLICACIÓN PERIÓDICA DE CARÁCTER SEMESTRAL DE LA UNIVERSIDAD FRANCISCO GAVIDIA
AÑO 26, n.º 63, ENERO-JUNIO 2026. SAN SALVADOR, EL SALVADOR, CENTROAMÉRICA

REALITY AND REFLECTION IS A BIENNIAL PERIODICAL PUBLICATION OF THE FRANCISCO GAVIDIA UNIVERSITY
YEAR 26, n.º 63, JANUARY-JUNE 2026. SAN SALVADOR, EL SALVADOR, CENTRAL AMERICA

Prácticas de evaluación auténtica en el aprendizaje del cálculo diferencial¹

Authentic assessment practices in learning differential calculus

Ismael Zamora Tovar

Licenciatura en Pedagogía, Universidad Autónoma de Guadalajara, México
Maestría en Educación, Universidad Autónoma de Guadalajara, México
Doctorado en Educación, Universidad Autónoma de Guadalajara, México
Coordinador del Modelo Educativo, Universidad Autónoma de Guadalajara, México
izamora@edu.uag.mx
<https://orcid.org/0000-0002-8520-1295>

Gilberto Sepúlveda Cervantes

Licenciatura en Sociología, Universidad de Guadalajara, México
Maestría en Gestión y Políticas de la Educación Superior, Universidad de Guadalajara, México
Especialista en Investigación Educativa, Universidad Autónoma de Guadalajara, México
gilberto.sepulveda@edu.uag.mx
<https://orcid.org/0000-0001-9331-5236>

¹ Para la elaboración de este documento se utilizaron herramientas de inteligencia artificial generativa como ChatGPT y Gemini, para la verificación de estilo y la revisión lingüística.

José Iván Reyes Cabrera

Ingeniero Mecánico Electricista, Universidad de Guadalajara, México
Maestría en Ciencias en Ingeniería Eléctrica, Universidad de Guadalajara, México
Doctorado en Ciencias en Ingeniería Eléctrica, Centro de Investigación
y Estudios Avanzados, Instituto Politécnico Nacional, México
Director del Departamento de Ciencias Básicas, Universidad Autónoma de Guadalajara, México
jose.reyes@edu.uag.mx
<https://orcid.org/0000-0001-8630-8396>

Fecha de recepción: 19 de noviembre de 2025

Fecha de aprobación: 17 de marzo de 2026

DOI:



RESUMEN

El estudio analiza la evaluación del aprendizaje en cálculo diferencial desde el enfoque de evaluación auténtica, en un contexto universitario donde persisten prácticas tradicionales. La literatura subraya la necesidad de instrumentos que valoren el pensamiento crítico, la resolución de problemas y la transferencia del conocimiento, además de considerar factores demográficos que influyen en la equidad y la validez de los procesos evaluativos. La investigación empleó un enfoque cuantitativo, descriptivo y transversal, con una muestra de 116 estudiantes. Se aplicó un cuestionario de 21 ítems en escala Likert, validado por expertos y con alta consistencia interna ($\alpha = 0.97$). El análisis estadístico incluyó estadística descriptiva e inferencial para identificar diferencias entre grupos demográficos. Los resultados muestran una alineación moderada de las prácticas de evaluación con los principios de la evaluación auténtica ($M = 3.9$). Las puntuaciones más altas se ubicaron en validez y confiabilidad (4.0), así como en pensamiento crítico y resolución de problemas (4.1). Otras dimensiones, como aplicación práctica, tareas retadoras, diversificación de instrumentos y transferencia del conocimiento, se situaron entre 3.8 y 3.9, lo que evidencia avances con margen de mejora. Algunos ítems, como la simulación de contextos reales y la adaptación a estilos de aprendizaje, mostraron mayor variabilidad en las percepciones estudiantiles. En conjunto, los hallazgos indican que, aunque existen elementos alineados con la evaluación auténtica, aún es necesario fortalecer la contextualización de las tareas, ampliar la diversidad de instrumentos y promover una mayor integración interdisciplinaria para favorecer aprendizajes más profundos, pertinentes y transferibles.

Palabras clave: evaluación auténtica, cálculo diferencial, prácticas de evaluación, transferencia del aprendizaje.

ABSTRACT

This study examines the assessment of learning in differential calculus from the perspective of authentic assessment, within a university context where traditional practices persist. The literature highlights the need for assessment tools that evaluate critical thinking, problem-solving, and the transfer of knowledge, while also taking into account demographic factors that influence the equity and validity of assessment processes. The research employed a quantitative, descriptive, and cross-sectional approach, with a sample of 116 students. A 21-item Likert-scale questionnaire was administered, validated by experts and with high internal consistency ($\alpha = 0.97$). Statistical analysis included descriptive and inferential statistics to identify differences among demographic groups. The results show moderate alignment of assessment practices with the principles of authentic assessment ($M = 3.9$). The highest scores were in validity and reliability (4.0), as well as in critical thinking and problem-solving (4.1). Other dimensions, such as practical application, challenging tasks, diversification of instruments, and knowledge transfer, ranged between 3.8 and 3.9, indicating progress with room for improvement. Some items, such as the simulation of real-world contexts and adaptation to learning styles, showed greater variability in student perceptions. Overall, the findings suggest that, while there are elements consistent with authentic assessment, there is still a need to strengthen the contextualization of tasks, expand the range of assessment tools, and promote greater interdisciplinary integration to foster deeper, more relevant, and transferable learning.

Keywords: *authentic assessment, differential calculus, assessment practices, transfer of learning.*

Introducción

En el contexto de la educación superior, la evaluación del aprendizaje atraviesa una etapa crítica de transformación (Boud, 2020), marcada por la creciente incongruencia entre los métodos evaluativos tradicionales y las demandas de una formación universitaria orientada al desarrollo de competencias complejas. Aún persisten prácticas centradas en exámenes estandarizados y pruebas memorísticas que reflejan poco la capacidad del estudiante para aplicar conocimientos en situaciones reales (Stankov, 2024), resolver problemas interdisciplinarios o transferir saberes a entornos laborales y sociales. Esta desconexión limita no solo el aprendizaje profundo y su aplicabilidad, sino también la calidad formativa de la educación superior frente a los desafíos contemporáneos.

Superar esta crisis exige repensar la evaluación como una herramienta formativa, integradora y contextualizada, capaz de acompañar el proceso de aprendizaje en todas sus dimensiones. Ello implica diseñar instrumentos que valoren la reflexión crítica, la resolución de problemas, la creatividad y la colaboración, promoviendo una evaluación auténtica y significativa (Ajjawi *et al.*, 2024). Bajo este enfoque, la evaluación deja de ser un acto final de medición y control para convertirse en un proceso continuo que retroalimenta la enseñanza, fortalece la autonomía estudiantil y asegura la pertinencia social del conocimiento universitario.

La evaluación del aprendizaje es un proceso fundamental en la educación, pues permite obtener información válida y confiable para la acreditación y certificación de los aprendizajes esperados establecidos en los programas de estudio y, en consecuencia, para el cumplimiento del perfil de egreso. Revisar periódicamente la pertinencia y relevancia de actividades, tareas y exámenes utilizados por los docentes proporciona insumos valiosos para el fortalecimiento de la enseñanza (Ibarra-Sáiz *et al.*, 2023).

La evaluación auténtica representa un enfoque pedagógico que trasciende los exámenes tradicionales de opción múltiple o preguntas cortas. En lugar de medir únicamente la memorización, busca valorar la capacidad de los estudiantes para aplicar conocimientos y habilidades en contextos significativos y reales. Su propósito es desafiar a los estudiantes a demostrar lo aprendido mediante tareas que simulan escenarios del mundo profesional, lo cual fomenta el pensamiento crítico, la resolución de problemas y la creatividad, al tiempo que conecta el contenido académico con experiencias significativas de aprendizaje (Vallejo Ruiz y Molina Saorín, 2014).

Este enfoque incorpora diversas estrategias: proyectos, portafolios, presentaciones, debates y experimentos, que permiten una valoración integral de distintas dimensiones del aprendizaje. Un aspecto relevante es la transferencia del conocimiento, es decir, la capacidad del estudiante para aplicar lo aprendido en contextos nuevos, diversos y reales, demostrando

la conexión entre teoría y práctica. Además de ser una práctica retadora (Ashford-Rowe *et al.*, 2014; Gijbels y Dochy, 2006), la evaluación auténtica es considerada objetiva, válida y confiable para medir aprendizajes universitarios (Tristán López y Pedraza Corpus, 2017).

Entre sus beneficios destacan el incremento en la motivación estudiantil, al proponer actividades más significativas y desafiantes; el desarrollo de competencias como comunicación, trabajo colaborativo, investigación y presentación de información; y la posibilidad de ofrecer retroalimentación más personalizada y precisa. Finalmente, al estar alineada con situaciones reales, esta forma de evaluación contribuye a una mejor preparación profesional, facilitando la transferencia de conocimientos y habilidades a escenarios laborales concretos (Vallejo-Ruiz y Molina-Saorín, 2014).

Para que una evaluación sea auténtica, deben considerarse factores que aseguren su pertinencia (Villarroel y Bruna, 2019). Entre ellos destacan la adaptación al contexto y a la diversidad estudiantil, el equilibrio con otras formas de evaluación y la centralidad de la retroalimentación, que favorece la reflexión y la autonomía. En definitiva, una evaluación de calidad debe percibirse como rigurosa, confiable, relevante y atractiva (Ibarra-Sáiz y Rodríguez-Gómez, 2020).

Para este estudio, la evaluación auténtica es un enfoque pedagógico que busca evaluar la capacidad de los estudiantes para aplicar conocimientos y habilidades en situaciones reales, y se define por seis dimensiones: la aplicación práctica del conocimiento a través de la simulación de contextos profesionales; el fomento del pensamiento crítico y la resolución de problemas, que exigen justificar respuestas y crear soluciones; el aprendizaje significativo, que conecta el contenido con la vida real para aumentar la motivación; la diversificación de instrumentos de evaluación, que se adaptan a distintos estilos de aprendizaje; la validez y confiabilidad de los instrumentos, asegurando que miden las competencias correctas; y la promoción de la transferencia del conocimiento a contextos nuevos e interdisciplinarios.

Una revisión crítica de la literatura sobre la evaluación del aprendizaje en educación superior evidencia que la evaluación del aprendizaje ha evolucionado en las últimas décadas, pasando de enfoques tradicionales centrados en la memorización y la reproducción de contenidos hacia prácticas más complejas que buscan promover la aplicación del conocimiento, la resolución de problemas, el aprendizaje significativo y la transferencia de saberes. Esta revisión de la literatura examina cómo las prácticas de evaluación universitaria se relacionan con criterios como la diversificación de instrumentos, la validez y confiabilidad, la inclusión de tareas retadoras y la transferencia del conocimiento.

La literatura distingue entre la evaluación del aprendizaje (sumativa) y la evaluación para el aprendizaje (formativa), reconociendo que ambas coexisten en la educación superior,

aunque la segunda facilita el aprendizaje a través de la retroalimentación (Hernández, 2012). Sin embargo, la implementación de prácticas auténticas y formativas sigue siendo limitada; menos del 30 % de los docentes universitarios aplican más de tres tareas auténticas en sus cursos, lo que restringe la aplicación del conocimiento y la resolución de problemas en contextos reales (Rawlusk, 2018).

La autoevaluación y la evaluación entre pares han demostrado ser estrategias efectivas para fomentar la autonomía, el pensamiento crítico y la transferencia del conocimiento, aunque su adopción aún es incipiente (Ifenthaler *et al.*, 2023; Alshammari, 2024). El uso de entornos digitales ha facilitado la integración de estas prácticas, permitiendo una retroalimentación más inmediata y el seguimiento del comportamiento de aprendizaje, lo que se asocia con un mejor desempeño académico (Ifenthaler *et al.*, 2023).

La diversificación de instrumentos de evaluación, como portafolios, tareas colaborativas y evaluaciones prácticas, se ha incrementado, especialmente en contextos de educación a distancia y durante la pandemia (Chavez y Lamorinas, 2023). Esta variedad contribuye a una mayor validez y confiabilidad en la medición del aprendizaje, especialmente cuando se combinan métodos directos e indirectos (Taylor *et al.*, 2023). No obstante, persisten desafíos en la alineación de los instrumentos con los objetivos de aprendizaje y en la capacitación docente para su uso efectivo (Holmes, 2024).

Las tareas retadoras, que exigen análisis, síntesis y aplicación, son recomendadas para promover el aprendizaje significativo y la transferencia; sin embargo, su implementación sigue siendo escasa (Rawlusk, 2018). La participación de los estudiantes en la evaluación, mediante la autoevaluación y la coevaluación, incrementa el nivel de desafío y compromiso, favoreciendo el desarrollo de habilidades de autorregulación y aprendizaje profundo (Rutherford *et al.*, 2025).

En esta perspectiva, la literatura coincide en la necesidad de avanzar hacia prácticas de evaluación más auténticas, diversificadas y centradas en el estudiante, que promuevan la aplicación del conocimiento, la resolución de problemas y la transferencia. Sin embargo, la adopción de estas prácticas enfrenta barreras institucionales y culturales, lo que limita su impacto en la educación universitaria.

Respecto a los factores demográficos asociados a las prácticas de evaluación del aprendizaje en estudiantes universitarios, la revisión de la literatura permite realizar las siguientes afirmaciones:

La equidad y la validez en la evaluación del aprendizaje universitario constituyen temas de amplio interés en la educación superior contemporánea. Diversos estudios han comenzado a explorar cómo factores demográficos, como género, etnia, nivel socioeconómico, edad y

estilos de aprendizaje, influyen en la aplicación del conocimiento, la resolución de problemas, el aprendizaje significativo, la diversificación de instrumentos de evaluación, la validez, la confiabilidad, las tareas retadoras y la transferencia del conocimiento.

En primer lugar, la literatura evidencia que los factores demográficos pueden afectar la imparcialidad y la validez de las prácticas de evaluación. Un estudio realizado en Irán identificó que el género, la etnia y el nivel socioeconómico (SES) pueden contaminar los resultados de la evaluación, generando desigualdades en la aplicación del conocimiento y en el acceso a tareas retadoras y transferibles. Los docentes reconocen que estos sesgos pueden tener efectos duraderos en la trayectoria académica de los estudiantes, comprometiendo la equidad y la confiabilidad de la evaluación (Rezai *et al.*, 2022). Respecto al sexo, las mujeres valoran más los aspectos colaborativos y psicosociales de la evaluación, mientras que los hombres muestran mayor precisión en la autoevaluación (González-Betancor *et al.*, 2017; Mortadi *et al.*, 2020). También se observan diferencias en la satisfacción general y en el aprovechamiento competencial (Morera-Castro *et al.*, 2023).

Por otro lado, los estilos de aprendizaje, asociados a variables personales y educativas, también influyen en el rendimiento según el tipo de instrumento de evaluación. Por ejemplo, los estudiantes de Psicología tienden a preferir estilos teóricos y abstractos, mientras que los de Educación Primaria muestran estilos más prácticos. Estas diferencias pueden llevar a que ciertos grupos demográficos se beneficien más de determinados métodos de evaluación, lo que subraya la importancia de diversificar los instrumentos para garantizar la validez y la equidad (Maya *et al.*, 2021).

La edad es otro factor relevante, ya que se asocia con cambios en los estilos de aprendizaje. Los estudiantes mayores tienden a preferir estilos reflexivos y globales, lo que implica la necesidad de adaptar periódicamente la instrucción y los instrumentos de evaluación para mantener la relevancia y el desafío de las tareas, favoreciendo así el aprendizaje significativo y la transferencia (Mehari *et al.*, 2024). En cuanto al rendimiento académico, los estudiantes con mejor promedio perciben mayor alineación entre las evaluaciones y las competencias profesionales, y demuestran mejor desempeño en evaluaciones entre pares (Mortadi *et al.*, 2020; Ravi y Besharat, 2026).

Las competencias digitales, necesarias para la diversificación de instrumentos y la aplicación del conocimiento en entornos virtuales, también presentan diferencias asociadas al género y al rendimiento académico previo. Estas brechas pueden influir en la capacidad de los estudiantes para afrontar tareas retadoras y transferir conocimientos en contextos digitales, aunque el tipo de institución de procedencia tiene menor impacto (Glasserman-Morales *et al.*, 2024).

Considerando la disciplina académica, la percepción de autenticidad varía según la carrera. Los estudiantes de Educación valoran más las metodologías activas, mientras que en áreas sanitarias y matemáticas se observan limitaciones en creatividad y aplicación práctica (Gil-Galván *et al.*, 2020; Ravi y Besharat, 2026).

La satisfacción general depende de la coherencia entre objetivos de aprendizaje, retroalimentación y pertinencia de las actividades evaluativas. Cuando se percibe autenticidad, aumentan la motivación y la implicación; cuando no, se genera insatisfacción (Morera-Castro *et al.*, 2023; Zueck *et al.*, 2023; Caba, 2023).

Como puede inferirse, aunque los factores demográficos muestran correlaciones pequeñas con el rendimiento académico global, su impacto puede ser relevante en la equidad, validez y confiabilidad de la evaluación, especialmente cuando interactúan con otros factores, como la autoeficacia y las estrategias de aprendizaje (Richardson *et al.*, 2012).

La literatura sugiere que los factores demográficos pueden influir en la equidad, validez y efectividad de las prácticas de evaluación universitaria, afectando la aplicación del conocimiento, la resolución de problemas y la transferencia. La diversificación de instrumentos y la adaptación a las características demográficas favorecen una evaluación más justa y significativa. La forma en que los estudiantes perciben la evaluación auténtica está influenciada por factores como el género, el rendimiento académico, la disciplina de estudio y la satisfacción general. Considerar estas variables resulta pertinente para crear evaluaciones justas y significativas que contribuyan al desarrollo integral de competencias profesionales.

En el caso específico de las matemáticas universitarias, la implementación de la evaluación auténtica busca posicionar a los estudiantes como actores centrales del conocimiento matemático a través de la solución de problemas significativos (Boaler, 2016). Los procesos que sustentan este enfoque incluyen la resolución de problemas complejos, el razonamiento para la demostración matemática y la conexión y representación de conceptos matemáticos (Schoenfeld, 2016; Thompson *et al.*, 2021). Esto contrasta con los métodos tradicionales, que enfatizan problemas cerrados con respuestas únicas, mientras que la evaluación auténtica prioriza el proceso de razonamiento sobre el resultado y el contexto significativo sobre los problemas artificiales (Rasmussen y Kwon, 2007).

Esta perspectiva adquiere especial relevancia en la asignatura de cálculo diferencial, donde el objetivo trasciende la ejecución procedimental para introducir al estudiante en conceptos matemáticos abstractos y en su comprensión conceptual profunda. Esto le permite comprender fenómenos dinámicos —como el movimiento, el crecimiento o el comportamiento de sistemas complejos— y le proporciona herramientas para analizar el cambio de una variable respecto a otra, con el propósito de optimizar recursos, predecir

comportamientos y controlar sistemas. Su carácter transversal en la ingeniería contribuye al desarrollo de competencias como resolución de problemas, argumentación lógica y pensamiento analítico, necesarias para la práctica profesional y la ciudadanía crítica (Cervantes Zubirías *et al.*, 2024).

Con base en estas consideraciones, la presente investigación tiene por objetivos: 1) Evaluar las prácticas de evaluación empleadas en la asignatura de cálculo diferencial en relación con los principios de la evaluación auténtica en el nivel universitario; y 2) Identificar los factores asociados a partir de los datos demográficos de la población estudiada.

El propósito es generar recomendaciones que contribuyan al fortalecimiento de las prácticas evaluativas bajo el enfoque de evaluación auténtica.

Hipótesis:

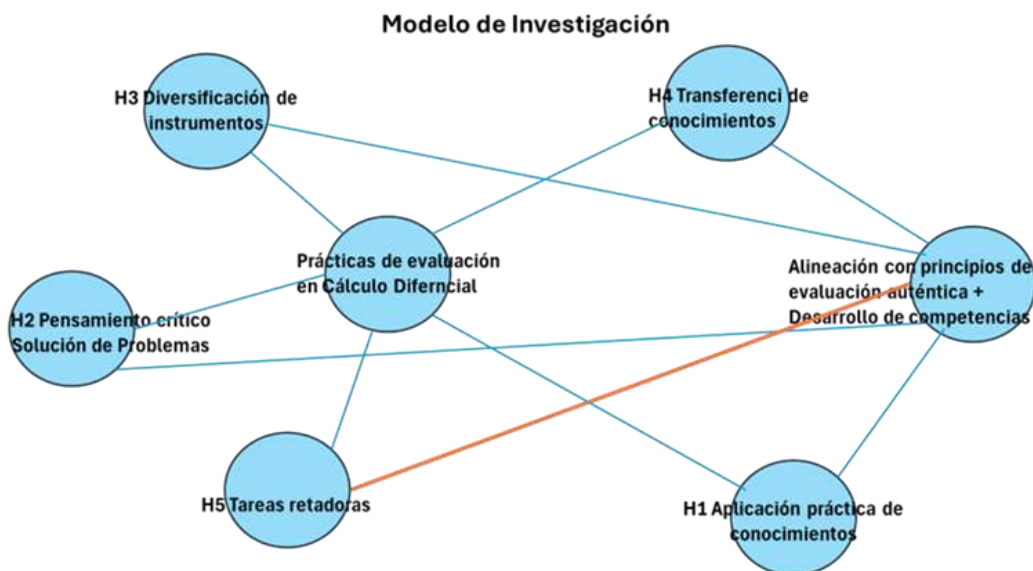
- HG: las prácticas de evaluación en cálculo diferencial no se alinean con los principios de evaluación auténtica.
- H1: las prácticas de evaluación se centran únicamente en la reproducción de contenidos, con escasa conexión con situaciones del mundo real.
- H2: los estudiantes perciben que las evaluaciones no fomentan el desarrollo del pensamiento crítico ni la resolución de problemas complejos.
- H3: no existe diversificación en los instrumentos de evaluación que permita atender distintos estilos de aprendizaje.
- H4: las evaluaciones no promueven la transferencia de conocimientos a contextos nuevos o interdisciplinarios.
- H5: no existe relación significativa entre las tareas retadoras, la aplicación práctica y la transferencia del conocimiento.

El modelo de investigación plantea que las prácticas de evaluación del aprendizaje en cálculo diferencial actúan como variable independiente y se relacionan con cinco dimensiones de la evaluación auténtica: aplicación práctica del conocimiento (H1), pensamiento crítico y resolución de problemas (H2), diversificación de instrumentos de evaluación (H3), transferencia del conocimiento (H4) y tareas retadoras (H5). Estas dimensiones, a su vez, inciden en la alineación con los principios de la evaluación auténtica y en el desarrollo de competencias transferibles, pensamiento crítico y aprendizaje significativo, considerados como variables dependientes.

El modelo enfatiza que todas las dimensiones están interconectadas, pero otorga especial relevancia a las tareas retadoras, las cuales muestran una relación más fuerte con la aplicación y transferencia del conocimiento, consolidándose como un factor clave en la mejora del aprendizaje auténtico (ver Figura 1)

Figura 1

Modelo de investigación



Nota. Elaboración propia.

Método

En este estudio cuantitativo, descriptivo y transversal se analizó la percepción de 116 estudiantes de cálculo diferencial sobre sus prácticas de evaluación, enfocándose en los principios de la evaluación auténtica. La investigación se realizó de enero a junio de 2025, garantizando la participación voluntaria y anónima de los estudiantes del Departamento de Ciencias Básicas.

Para recabar la información, se diseñó una encuesta de 21 preguntas con una escala tipo Likert de cinco puntos, que abarcó seis dimensiones de la evaluación auténtica. Este instrumento fue revisado y aprobado por expertos en el área. La consistencia interna general de la encuesta fue muy alta (alfa de Cronbach = 0.97), y las dimensiones individuales también mostraron una fiabilidad considerable, con valores entre 0.74 y 0.91, lo que confirma su solidez y confiabilidad para el estudio.

Los datos se analizaron con el *software* estadístico Statgraphics. Se utilizó estadística descriptiva —medias, desviaciones estándar y frecuencias— para resumir las percepciones de los estudiantes, y estadística inferencial —pruebas t de Student o ANOVA— para identificar diferencias significativas entre grupos demográficos. Este método permitió realizar un análisis riguroso de las prácticas de evaluación auténtica en la universidad.

Resultados

Para medir la alineación de las prácticas docentes con la evaluación auténtica, se utilizó el siguiente criterio: sin alineación, $\text{media} \leq 3.0$; alineación moderada, $3.0 < \text{media} < 4.0$; y alineación alta, $\text{media} \geq 4.0$. Con referencia a los objetivos de investigación:

Para el objetivo 1, sobre la evaluación auténtica

En la Tabla 1 se muestra cómo los estudiantes perciben que las prácticas de evaluación en la asignatura de cálculo diferencial están moderadamente alineadas con los principios de la evaluación auténtica. La puntuación promedio general de 3.9, en una escala de 1 a 5, y las bajas desviaciones estándar (DS) indican una percepción consistente y favorable.

Respecto a las puntuaciones altas (4.0 o más), los estudiantes perciben que las evaluaciones presentan una alta alineación con la dimensión de validez y confiabilidad (PROM = 4.0). Asimismo, consideran que existe una alta alineación con la dimensión «promueve el pensamiento crítico y la resolución de problemas» (PROM = 4.1).

En relación con las puntuaciones moderadas (3.6 a 3.9), la dimensión de tareas retadoras presenta una alineación moderada (PROM = 3.9), exigiendo cierto esfuerzo intelectual más allá de la memorización. Las dimensiones de aplicación práctica del conocimiento (PROM = 3.9), aprendizaje significativo (PROM = 3.8), diversificación de instrumentos (PROM = 3.8) y transferencia del conocimiento (PROM = 3.9) obtuvieron puntuaciones similares. Esto sugiere que, aunque los estudiantes consideran que estos aspectos están presentes, existe margen de mejora para fortalecer la conexión de los contenidos con el mundo real, la variedad de las actividades y la integración de saberes.

Por otra parte, las desviaciones estándar indican que la mayoría de estas son bajas (inferiores a 1.1), lo que refleja un acuerdo generalizado entre los estudiantes. Sin embargo, los ítems de «simulación de situaciones del mundo real» (DS = 1.07) y «tareas que se adaptan a distintos estilos de aprendizaje» (DS = 1.20) presentan una dispersión más alta, lo que sugiere una mayor variación de opiniones entre los encuestados respecto a estos puntos específicos.

Tabla 1

Evaluación auténtica-cálculo diferencial

Evaluación auténtica	PROM	DS
	3.9	0.77
Aplicación práctica del conocimiento	3.9	0.80
1. Las tareas solicitadas a los estudiantes simulan situaciones del mundo real relacionadas con la asignatura.	3.6	1.07

Evaluación auténtica	PROM	DS
	3.9	0.77
2. Las actividades de evaluación permiten aplicar conocimientos en contextos profesionales auténticos.	3.8	1.01
3. Los exámenes u otras evaluaciones de la asignatura requieren más que la memorización de contenidos, exigiendo análisis o síntesis.	4.2	0.86
Pensamiento crítico y resolución de problemas	4.1	0.77
4. Las evaluaciones promueven el análisis de información compleja.	4.1	0.86
5. Se fomenta que los estudiantes justifiquen sus respuestas y decisiones.	4.2	0.98
6. Las actividades permiten desarrollar y demostrar pensamiento crítico y creativo.	4.1	0.94
Aprendizaje significativo	3.8	1.00
7. Las evaluaciones están diseñadas para conectar el contenido del curso con situaciones relevantes de la vida real.	3.7	1.09
8. Los estudiantes pueden ver claramente la utilidad de lo que están aprendiendo.	3.8	1.07
9. Las actividades de evaluación fortalecen la motivación y el interés por la materia.	3.8	1.10
Diversificación de instrumentos de evaluación	3.8	0.90
10. Se emplean diversos tipos de evaluación (proyectos, debates, portafolios, etc.) para valorar el aprendizaje.	4.0	0.96
11. Se permite a los estudiantes demostrar sus conocimientos a través de diferentes medios.	3.8	0.99
12. Las tareas se adaptan a distintos estilos de aprendizaje y habilidades.	3.7	1.20
Validez y confiabilidad de la evaluación	4.0	0.88
13. Los instrumentos utilizados para evaluar están alineados con los aprendizajes esperados del programa.	4.1	0.87
14. Las evaluaciones aplicadas permiten medir de manera válida las competencias del perfil de egreso.	3.9	1.01
15. Las calificaciones obtenidas reflejan de forma confiable el nivel de aprendizaje del estudiante.	3.8	1.08
Tareas retadoras	3.9	0.79
16. Las tareas asignadas exigen un nivel de esfuerzo intelectual elevado, más allá de repetir o memorizar información.	4.0	0.80
17. Las actividades requieren que el estudiante busque, seleccione y utilice información de diferentes fuentes bibliográficas.	3.9	0.94
18. Las evaluaciones incluyen situaciones complejas que demandan soluciones creativas e innovadoras.	3.7	1.04
Transferencia del conocimiento	3.9	0.92
19. Las evaluaciones permiten aplicar conocimientos y habilidades a contextos nuevos o distintos al estudiado originalmente.	3.8	1.04
20. Se fomenta que los estudiantes integren saberes de diferentes asignaturas o experiencias previas.	3.9	0.99

Evaluación auténtica	PROM	DS
	3.9	0.77
21. Las tareas promueven el uso del conocimiento adquirido para resolver problemas reales fuera del aula.	3.9	1.08

Nota. Elaboración propia.

La Figura 2 muestra los resultados globales de la evaluación de las prácticas de evaluación del aprendizaje de cálculo diferencial. Se presenta la media (PROM) y la desviación estándar (DS) de la percepción de los estudiantes en las diferentes dimensiones de la evaluación. La puntuación global para la evaluación auténtica es de 3.9, con una desviación estándar de 0.77. Esto indica que, en general, los estudiantes perciben las prácticas de evaluación como moderadamente alineadas con los principios de la evaluación auténtica.

La dimensión de validez y confiabilidad de la evaluación, con una media de 4.0 y una DS de 0.88, presenta la percepción más alta. Los estudiantes consideran que los instrumentos de evaluación son válidos y confiables, lo que significa que están bien alineados con los objetivos de aprendizaje y que las calificaciones reflejan de forma precisa el nivel de conocimiento.

La dimensión de tareas retadoras también presenta una media alta de 4.0 (DS = 0.88). Los estudiantes perciben que las tareas asignadas son desafiantes y requieren un nivel de esfuerzo intelectual significativo, lo que va más allá de la simple memorización.

Respecto al pensamiento crítico y resolución de problemas, con una media de 4.1 y una DS de 0.77, los estudiantes consideran que las evaluaciones fomentan el análisis y la justificación de respuestas, promoviendo el pensamiento crítico. Esta dimensión, junto con las dos anteriores, presenta las medias más altas, lo que indica que, en estos aspectos, las prácticas de evaluación están muy bien implementadas.

La aplicación práctica del conocimiento, con una media de 3.9 (DS = 0.80), sugiere que las tareas de evaluación permiten aplicar los conocimientos en contextos del mundo real y profesional, aunque existe margen de mejora.

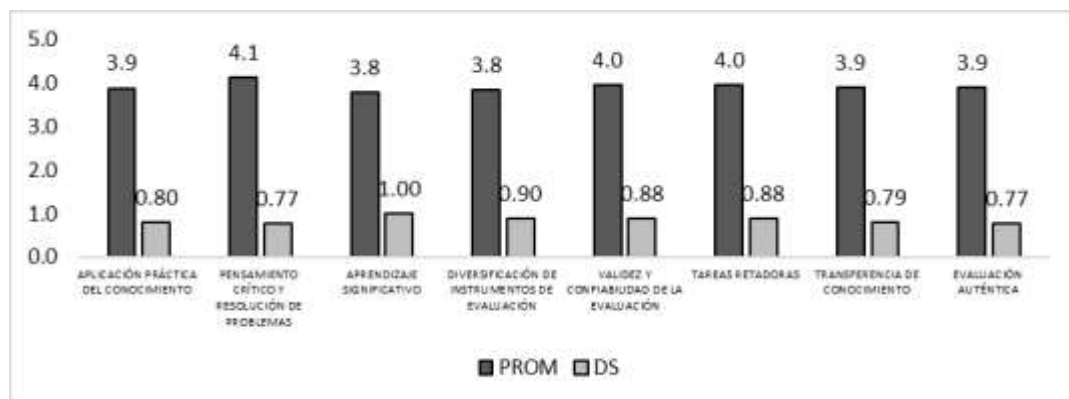
La transferencia del conocimiento, con una media de 3.9 y una DS de 0.79, muestra que los estudiantes perciben que las evaluaciones promueven la aplicación de conocimientos a nuevos contextos y la integración de saberes, aunque no de manera sobresaliente.

El aprendizaje significativo, con una media de 3.8 (DS = 1.00), indica que las evaluaciones contribuyen a que los estudiantes reconozcan la utilidad de lo que aprenden y fortalezcan su motivación; sin embargo, esta es la dimensión con mayor variabilidad en las respuestas, al presentar la desviación estándar más alta.

Respecto a la diversificación de instrumentos de evaluación, con una media de 3.8 (DS = 0.90), los estudiantes perciben que, aunque se utilizan diferentes tipos de evaluaciones, esta práctica podría mejorar para adaptarse de mejor manera a distintos estilos de aprendizaje y permitir demostrar habilidades de forma más integral.

Figura 2

Evaluación auténtica-cálculo diferencial (resumen)



Nota. Elaboración propia.

En general, las dimensiones con las puntuaciones más altas son aquellas relacionadas con la validez, el pensamiento crítico y el carácter retador de las tareas, mientras que las dimensiones con puntuaciones ligeramente más bajas, aunque todavía dentro de un rango favorable, corresponden a la diversificación de instrumentos y al aprendizaje significativo.

En síntesis, los resultados reflejan una valoración positiva de la evaluación auténtica, especialmente en el fomento del pensamiento crítico y en la validez de los instrumentos empleados. Aunque algunas dimensiones presentan puntuaciones menores, la consistencia observada en las respuestas confirma la pertinencia del modelo como estrategia para fortalecer el aprendizaje en la asignatura de cálculo diferencial.

Para el objetivo 2 sobre los factores asociados a la evaluación

La Tabla 2 de análisis de varianza (ANOVA) revela que ni el sexo ni el promedio de carrera tienen un efecto significativo en la percepción de los estudiantes sobre las prácticas de evaluación auténtica. Esto se confirma al observar los valores *p* para ambas variables, que corresponden a 0.4639 (sexo) y 0.2738 (promedio de carrera). Dado que ambos valores son superiores al umbral de significancia de 0.05, se concluye que no existen diferencias estadísticamente significativas entre los grupos. En otras palabras, la percepción sobre la

evaluación es consistente entre hombres y mujeres, así como entre estudiantes con promedios académicos altos, medios o bajos.

Tabla 2

Análisis de varianza para PROM GNRL - Suma de cuadrados tipo III

Fuente	Suma de cuadrados	Gl	Cuadrado medio	Razón-F	Valor-P
Efectos principales					
A:Sexo	0.328711	1	0.328711	0.54	0.4639
B:Prom Carr.	2.39656	3	0.798854	1.31	0.2738
Residuos	67.5349	111	0.608423		
Total (corregido)	70.0114	115			
Todas las razones-F se basan en el cuadrado medio del error residual					

Nota. Elaboración propia.

Respecto a las hipótesis

Para la **hipótesis general** referida en la Tabla 3, relacionada con la alineación de los principios de evaluación auténtica, se muestran a continuación las correlaciones producto-momento de Pearson entre cada par de variables incluidas en el estudio. Estos coeficientes indican la fuerza y dirección de la relación lineal, con valores que oscilan entre -1 y +1. Junto a cada coeficiente se incluye, entre paréntesis, el número de pares de datos empleados en el cálculo, así como el valor p correspondiente. Este último permite determinar la significancia estadística de las correlaciones estimadas, considerando que valores p inferiores a 0.05 reflejan relaciones significativamente diferentes de cero, con un nivel de confianza del 95 %.

En el análisis realizado se observa que todas las correlaciones entre las variables presentan valores p inferiores a 0.05, lo que indica que las relaciones encontradas son estadísticamente significativas. Este hallazgo sugiere una asociación consistente entre las dimensiones evaluadas, lo que respalda la idea de que las prácticas analizadas están interrelacionadas y, por tanto, influyen conjuntamente en la experiencia de evaluación auténtica en la asignatura.

Tabla 3

Correlaciones de Pearson entre dimensiones de evaluación auténtica

	AP- conocimiento	PCyR- problemas	Aprendizaje significativo	DI- evaluación	Tareas retadoras	T-conocimiento
Aplicación de conocimiento		0.6636 (116)	0.7663 (116)	0.7360 (116)	0.6990 (116)	0.7241 (116)
		0.0000	0.0000	0.0000	0.0000	0.0000
	0.6636		0.6736	0.6209	0.6943	0.6735

	AP- conocimiento	PCyR- problemas	Aprendizaje significativo	DI- evaluación	Tareas retadoras	T-conocimiento
PCyR- problemas	(116) 0.0000		(116) 0.0000	(116) 0.0000	(116) 0.0000	(116) 0.0000
Aprendizaje significativo	0.7663 (116) 0.0000	0.6736 (116) 0.0000		0.8472 (116) 0.0000	0.7354 (116) 0.0000	0.8353 (116) 0.0000
DI-evaluación	0.7360 (116) 0.0000	0.6209 (116) 0.0000	0.8472 (116) 0.0000		0.7940 (116) 0.0000	0.8580 (116) 0.0000
Tareas retadoras	0.6990 (116) 0.0000	0.6943 (116) 0.0000	0.7354 (116) 0.0000	0.7940 (116) 0.0000		0.8013 (116) 0.0000
Transferencia conocimiento	0.7241 (116) 0.0000	0.6735 (116) 0.0000	0.8353 (116) 0.0000	0.8580 (116) 0.0000	0.8013 (116) 0.0000	

Nota. Elaboración propia.

Los coeficientes obtenidos muestran relaciones fuertes y positivas entre todas las dimensiones, con valores que oscilan entre 0.62 y 0.86, lo que indica que, a medida que los estudiantes perciben un mayor uso de una práctica de evaluación, también perciben un mayor uso de las demás. La correlación más alta se encuentra entre la diversificación de instrumentos (DI-evaluación) y la transferencia del conocimiento (T-conocimiento) ($r = 0.86$), lo que sugiere que el uso de una variedad de herramientas evaluativas facilita la capacidad de los estudiantes para aplicar sus conocimientos en nuevos contextos. Además, el aprendizaje significativo (APR-significativo) y la transferencia del conocimiento están fuertemente relacionados ($r = 0.84$), lo que implica que, cuando el aprendizaje es percibido como relevante, es más probable que los estudiantes transfieran esos conocimientos a situaciones reales.

Bajo esta evidencia, se comprueba que las prácticas de evaluación en cálculo diferencial sí muestran una alineación moderada con los principios de evaluación auténtica. Esto deja abierta la posibilidad de establecer mejores prácticas que permitan alcanzar un nivel de alineación más significativo.

Para la **hipótesis 1**, en la Tabla 4 sobre resultados de la prueba t de Student, el análisis muestra que los estudiantes perciben un uso significativo de la aplicación práctica del conocimiento en sus evaluaciones, con una puntuación media de 3.88 en una escala de 5. Este resultado es estadísticamente significativo ($p \approx 0$), lo que confirma que la percepción no es producto del azar. Aunque la media es alta, se plantea la posibilidad de que el enfoque de las evaluaciones aún podría inclinarse hacia la reproducción de contenidos más que hacia la aplicación en el mundo real, a pesar de que los estudiantes valoran su contribución al aprendizaje.

Tabla 4

Prueba t de Student para aplicación práctica del conocimiento

Prueba de hipótesis para AP-conocimiento
Media muestral = 3.88103
Mediana muestral = 4.0
Desviación estándar de la muestra = 0.804002
Prueba t
Hipótesis nula: media = 3.0
Alternativa: no igual
Estadístico t = 11.8022
Valor-P = 0
Se rechaza la hipótesis nula para alfa = 0.05.

Nota. Elaboración propia.

Para la **hipótesis 2**, de acuerdo con la Tabla 5, los estudiantes percibieron que las evaluaciones de la asignatura promueven significativamente el pensamiento crítico y la resolución de problemas complejos, con una puntuación promedio de 4.12 en una escala de 5. El análisis estadístico, mediante la prueba t de Student, confirmó que esta percepción es significativamente mayor que el punto de referencia, indicando que las evaluaciones son consideradas efectivas en este aspecto. A pesar de estos resultados favorables, aún existe la posibilidad de mejorar el diseño de las tareas para fomentar un análisis más profundo y la resolución de problemas en contextos de mayor complejidad.

Tabla 5

Prueba T para pensamiento crítico y resolución de problemas

Prueba de hipótesis para PCYR-problemas
Media muestral = 4.12328
Mediana muestral = 4.3
Desviación estándar de la muestra = 0.776319
Prueba t
Hipótesis nula: media = 3.0
Alternativa: no igual
Estadístico t = 15.5839
Valor-P = 0
Se rechaza la hipótesis nula para alfa = 0.05.

Nota. Elaboración propia.

Para la **hipótesis 3**, de acuerdo con el análisis presentado en la Tabla 6, los estudiantes perciben un uso significativo de la diversificación de instrumentos de evaluación, con una puntuación media de 3.84 en una escala de 5. Este hallazgo es estadísticamente significativo ($p \approx 0$), lo que permite rechazar la hipótesis nula y sugiere que la variedad de herramientas de evaluación no es resultado del azar. Aunque los estudiantes reconocen el esfuerzo por incorporar diversos métodos, el texto señala la necesidad de continuar mejorando la flexibilidad e inclusión de estas estrategias para medir de manera más completa las competencias desarrolladas.

Tabla 6

Prueba t de Student para diversificación de instrumentos de evaluación

Prueba de hipótesis para DI-evaluación
Media muestral = 3.84138
Mediana muestral = 4.0
Desviación estándar de la muestra = 0.901117
Prueba t
Hipótesis nula: media = 3.0
Alternativa: no igual
Estadístico t = 15.5839
Valor-P = 3.8197E-7
Se rechaza la hipótesis nula para alfa = 0.05.

Nota. Elaboración propia.

Para la **hipótesis 4**, en la Tabla 7, los resultados de la prueba t de Student demuestran que la percepción de los estudiantes sobre la transferencia del conocimiento es significativamente mayor que el valor de referencia, lo que indica que las evaluaciones sí promueven esta habilidad. La media muestral de 3.86 y el valor p extremadamente bajo (3.4172×10^{-7}) confirman que este resultado es estadísticamente significativo y no se debe al azar. Al rechazarse la hipótesis nula, se valida que la diferencia entre la percepción de los estudiantes y el valor hipotético es real y que las prácticas de evaluación en la asignatura contribuyen de manera efectiva a la capacidad de los estudiantes para aplicar y transferir sus conocimientos a nuevos contextos.

Tabla 7

Prueba t de Student para transferencia del conocimiento

Prueba de hipótesis T-conocimiento
Media muestral = 3.85776
Mediana muestral = 4.0

Desviación estándar de la muestra = 0.930745

Prueba t

Hipótesis nula: media = 3.0

Alternativa: no igual

Estadístico t = 9.92575

Valor-P = **3.4172E-7**

Se rechaza la hipótesis nula para alfa = 0.05.

Nota. Elaboración propia.

Para la **hipótesis 5**, el análisis de correlación de Pearson, mostrado en la Tabla 8, confirma que existe una relación fuerte y positiva entre las tareas retadoras y la capacidad de los estudiantes para aplicar y transferir conocimientos. Los resultados muestran que las tareas que los estudiantes perciben como más desafiantes se asocian directamente con una mayor habilidad para transferir lo aprendido a nuevos contextos ($r = 0.8013$) y con la aplicación práctica del conocimiento ($r = 0.6990$). Dado que los valores p son 0.0000, estas correlaciones son estadísticamente significativas, lo que indica que no se deben al azar. En síntesis, el estudio sugiere que las tareas complejas y exigentes están fuertemente asociadas con el desarrollo de competencias de aplicación y transferencia del conocimiento.

Tabla 8

Correlaciones de Pearson entre dimensiones de tareas retadoras y su efecto en la aplicación y transferencia del conocimiento

	T-retadoras	AP-conocimiento	T-conocimiento
Tareas retadoras		0.6990	0.8013
		(116)	(116)
		0.0000	0.0000
Aplicación de conocimiento	0.6990		0.7241
	(116)		(116)
	0.0000		0.0000
Transferencia del conocimiento	0.8013	0.7241	
	(116)	(116)	
	0.0000	0.0000	

Nota. Elaboración propia.

Los resultados respaldan firmemente la **hipótesis 5**, concluyendo que la incorporación de tareas percibidas como retadoras constituye un factor pedagógico de gran relevancia. No solo fomenta de manera individual la aplicación y la transferencia de conocimientos, sino que estas competencias se refuerzan mutuamente dentro de un sistema interdependiente. Esto subraya que el desafío académico, lejos de representar una barrera, actúa como un

catalizador para el desarrollo de un aprendizaje profundo, significativo y funcional en los estudiantes.

En conjunto, los resultados ofrecen una visión consistente sobre la forma en que los estudiantes perciben las prácticas de evaluación en cálculo diferencial y muestran patrones claros tanto en la alineación con los principios de evaluación auténtica como en las relaciones entre sus distintas dimensiones. La evidencia cuantitativa respalda que, aunque existen fortalezas consolidadas, particularmente en pensamiento crítico, validez de los instrumentos y tareas retadoras, también persisten áreas con potencial de mejora relacionadas con la diversificación y la conexión con contextos reales. Asimismo, las asociaciones estadísticas identificadas y la ausencia de diferencias entre grupos permiten delinear un panorama integral sobre el comportamiento de las prácticas evaluativas. Estos elementos, en conjunto, abren la posibilidad de examinar con mayor profundidad el significado pedagógico de los hallazgos y sus implicaciones, aspecto que se desarrolla en el siguiente apartado.

Discusión

A partir de los hallazgos del estudio, las conclusiones destacan la valoración positiva que los estudiantes otorgan a la evaluación auténtica y la forma en que sus dimensiones se relacionan entre sí. En términos generales, se observó que las prácticas de evaluación en la asignatura de cálculo diferencial presentan una alineación moderada con los principios de la evaluación auténtica, lo cual se refleja en una percepción favorable, con un promedio de 3.9 en una escala de 5.

Respecto a las hipótesis, los resultados no respaldaron la hipótesis general ni las específicas H1-H4, pues la evidencia contradice la idea de que las prácticas de evaluación se limiten a la simple reproducción de contenidos o carezcan de promoción del pensamiento crítico, la diversificación de instrumentos y la transferencia de conocimientos. Por el contrario, las pruebas estadísticas, mediante *t* de Student, evidenciaron que la valoración de los estudiantes en estas dimensiones es significativamente superior al punto de referencia, lo que confirma una apreciación positiva de las prácticas implementadas. En contraste, la hipótesis 5 sí fue confirmada, al encontrarse una relación sólida y positiva entre las tareas retadoras y la capacidad de aplicar y transferir aprendizajes, lo que indica que, cuanto más desafiantes perciben los estudiantes las actividades, mayor es su habilidad para utilizar el conocimiento en contextos nuevos.

En cuanto a factores adicionales, el análisis de varianza (ANOVA) mostró que ni el sexo ni el promedio académico influyen significativamente en la percepción de la evaluación auténtica, lo que refleja una opinión homogénea entre los distintos grupos. Finalmente, al analizar la interconexión de las dimensiones, los resultados de la correlación de Pearson

muestran una fuerte y positiva interrelación entre todas las dimensiones de la evaluación auténtica. **Esto sugiere que la evaluación auténtica constituye un constructo coherente en el que cada una de sus partes se refuerza mutuamente**, impactando de manera conjunta en la experiencia de evaluación de los estudiantes.

Los resultados de este estudio muestran que las prácticas de evaluación en cálculo diferencial se perciben moderadamente alineadas con los principios de la evaluación auténtica ($M = 3.9/5$), lo que coincide con investigaciones que destacan su potencial para promover motivación, pensamiento crítico y transferencia de aprendizajes (Ajjawi *et al.*, 2024; Vallejo Ruiz y Molina Saorín, 2014). La alta valoración de las dimensiones de pensamiento crítico, validez de los instrumentos y tareas retadoras confirma lo planteado por Ashford-Rowe *et al.* (2014) y Rutherford *et al.* (2025), al señalar que los desafíos académicos y la retroalimentación formativa fortalecen el análisis y la aplicación del conocimiento en nuevos contextos.

Sin embargo, dimensiones como la diversificación de instrumentos y el aprendizaje significativo obtuvieron puntuaciones más bajas, lo que respalda las observaciones de Holmes (2024) y Villarroel y Bruna (2019) sobre las dificultades para integrar plenamente estrategias variadas y pertinentes en la práctica docente.

En contraste con estudios que identifican sesgos demográficos en la percepción de la evaluación (González-Betancor *et al.*, 2017; Rezai *et al.*, 2022), este trabajo no encontró diferencias significativas por sexo ni promedio académico, lo que sugiere homogeneidad en las opiniones estudiantiles. Asimismo, la correlación positiva entre todas las dimensiones confirma lo señalado por Gijbels y Dochy (2006) y Gulikers *et al.* (2008) respecto a la coherencia interna de la evaluación auténtica como constructo pedagógico. En conjunto, estos hallazgos respaldan la necesidad de consolidar prácticas evaluativas que no solo midan contenidos, sino que también promuevan la transferencia y el aprendizaje significativo, en consonancia con lo propuesto por Boud (2020).

Finalmente, en el contexto de la investigación en educación matemática universitaria, la confirmación de la hipótesis 5 concuerda con lo encontrado por Kogan y Laursen (2024), quienes señalan que la exposición a tareas complejas y desafiantes permitió que los estudiantes mostraran mejoras en cursos posteriores. De manera similar, Rasmussen y Kwon (2007) reportan que los enfoques centrados en problemas auténticos generan mejoras significativas en la comprensión conceptual y en las habilidades de razonamiento matemático aplicado. Esto demuestra que, en el contexto de la enseñanza de las matemáticas, el desafío cognitivo constituye un factor determinante para el desarrollo de competencias transferibles.

Respecto a futuras investigaciones, se sugiere ampliar el alcance del estudio hacia otras asignaturas y contextos institucionales, con el propósito de contrastar si la percepción sobre la evaluación auténtica se mantiene en diferentes escenarios. Asimismo, sería pertinente profundizar en la relación entre tareas retadoras y transferencia de conocimientos mediante estudios longitudinales, así como explorar el papel del docente en la implementación de estas prácticas y su impacto en la percepción estudiantil. Finalmente, se propone vincular la evaluación auténtica con variables de rendimiento académico y factores psicosociales, lo que permitiría comprender de manera más integral sus efectos en la formación universitaria.

Referencias

- Ajjawi, R., Tai, J., Dollinger, M., Dawson, P., Boud, D. y Bearman, M. (2024). From authentic assessment to authenticity in assessment: Broadening perspectives. *Assessment & Evaluation in Higher Education*, 49(4), pp. 499–510. <https://doi.org/10.1080/02602938.2023.2271193>
- Alshammari, A. (2024). Promoting Saudi university students' experience through peer assessment practice. *Theory and Practice in Language Studies*, 14(8), pp. 2484–2490. <https://doi.org/10.17507/tpsls.1408.21>
- Ashford-Rowe, K., Herrington, J. y Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education*, 39(2), pp. 205–222. <https://doi.org/10.1080/02602938.2013.819566>
- Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching*. Jossey-Bass/Wiley.
- Boud, D. (2020). Retos en la reforma de la evaluación en educación superior: Una mirada desde la lejanía. *RELIEVE*, 26(1), art. M3. <http://doi.org/10.7203/relieve.26.1.17088>
- Caba Liriano, G. M. (2023). Percepción de los estudiantes universitarios ante una estrategia de autoevaluación. *Cuaderno de Pedagogía Universitaria*, 20(39), pp. 177–187. <https://doi.org/10.29197/cpu.v20i39.495>
- Cervantes Zubirías, G., Morales Rodríguez, M. A., Vargas Orozco, C. M., Hernández Rodríguez, I. y Camacho Sánchez, F. Y. (2024). Estrategias lúdicas para el aprendizaje de cálculo diferencial. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, 5(4), pp. 2249–2260. <https://doi.org/10.56712/latam.v5i4.2410>
- Chavez, J. V. y Lamorinas, D. D. (2023). Reconfiguring assessment practices and strategies in online education during the pandemic. *International Journal of Assessment Tools in Education*, 10(1), pp. 160–174. <https://doi.org/10.21449/ijate.1094589>
- Gijbels, D. y Dochy, F. (2006). Students' assessment preferences and approaches to learning: can formative assessment make a difference? *Educational Studies*, 32(4), pp. 399–409. <https://doi.org/10.1080/03055690600850354>
- Gil-Galván, R., Martín-Espinosa, I. y Gil-Galvan, F. J. (2020). Percepciones de los estudiantes universitarios sobre las competencias adquiridas mediante el aprendizaje basado en problemas. *Educación XX1*, 24(1). <https://doi.org/10.5944/educxx1.26800>

- Glasserman-Morales, L. D., Alcantar-Nieblas, C. y Sisto, M. I. (2024). Demographic and school factors associated with digital competences in higher education students. *Contemporary Educational Technology*, 16(2), ep498. <https://doi.org/10.30935/cedtech/14288>
- González-Betancor, S. M., Bolívar-Cruz, A. y Verano-Tacoronte, D. (2017). Self-assessment accuracy in higher education: The influence of gender and performance of university students. *Active Learning in Higher Education*, 20(2), pp. 101–114. <https://doi.org/10.1177/1469787417735604>
- Gulikers, J. T. M., Kester, L., Kirschner, P. A. y Bastiaens, T. J. (2008). The effect of practical experience on perceptions of assessment authenticity, study approach, and learning outcome. *Learning and Instruction*, 18(2), pp. 172–186. <https://doi.org/10.1016/j.learninstruc.2007.02.012>
- Hernández, R. (2012). Does continuous assessment in higher education support student learning? *Higher Education*, 64(4), pp. 489–502. <https://doi.org/10.1007/s10734-012-9506-7>
- Holmes, AGD (2024). Desarrollo de la alfabetización en evaluación de estudiantes universitarios: teoría y práctica. *Innovare Journal of Education* , 12 (5), pp. 26-31. <https://doi.org/10.22159/ijoe.2024v12i5.52282>
- Ibarra-Sáiz, M. S. y Rodríguez-Gómez, G. (2020). Aprendiendo a evaluar para aprender en la educación superior. *Revista Iberoamericana de Evaluación Educativa*, 13(1), pp. 5–8. <https://revistas.uam.es/rie/article/view/12070>
- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., Lukas-Mujika, J. F. y Santos-Berrondo, A. (2023). Medios e instrumentos para evaluar los resultados de aprendizaje en másteres universitarios: Análisis de la percepción del profesorado sobre su práctica evaluativa. *Educación XX1*, 26(1), pp. 21–45. <https://doi.org/10.5944/educxx1.33443>
- Ifenthaler, D., Schumacher, C. y Kuzilek, J. (2023). Investigating students' use of self-assessments in higher education using learning analytics. *Journal of Computer Assisted Learning*, 39(1), pp. 255–268. <https://doi.org/10.1111/jcal.12744>
- Kogan, M. y Laursen, S. L. (2014). Assessing long-term effects of inquiry-based learning: A case study from college mathematics. *Innovative Higher Education*, 39(3), pp. 183–199. <https://doi.org/10.1007/s10755-013-9269-9>
- Maya, J., Luesia, J. F., & Pérez-Padilla, J. (2021). The Relationship between Learning Styles and Academic Performance: Consistency among Multiple Assessment Methods in Psychology and Education Students. *Sustainability*, 13(6), 3341. <https://doi.org/10.3390/su13063341>
- Mehari, A., Kassahun, B., Berhanu, H., & Birru, B. (2024). Describing learning styles and associated demographic characteristics among university students: Implications for instructional provision in a higher education context. *Kotebe Journal of Education*, 2(2), pp. 22–43. <https://www.bohrium.com/en/paper-details/describing-learning-styles-and-associated-demographic-characteristics-among-university-students-implications-for-instructional-provision-in-a-higher-education-context/1099914032552345657-110197>
- Morera-Castro, M., Nieva-Boza, C., Anchía-Umaña, I. y Herrera-González, E. (2023). Perception of Formative Competency Assessment and Virtual Course Satisfaction by Sex and Career

- Level among College Students. *Revista Electrónica Educare*, 27(3), pp. 1-22. <https://doi.org/10.15359/ree.27-3.17964>
- Mortadi, N. A., Al-Houry, S. S., Alzoubi, K. H., & Khabour, O. F. (2020). Effectiveness of peer evaluation in learning process: A case from dental technology students. *The Open Dentistry Journal*, 14, pp. 692–697. <https://doi.org/10.2174/1874210602014010692>
- Rasmussen, C. y Kwon, O. N. (2007). An inquiry-oriented approach to undergraduate mathematics. *Journal of Mathematical Behavior*, 26(3), pp. 189-194. <https://doi.org/10.1016/j.jmathb.2007.10.001>
- Ravi, M. y Besharat, M. (2026). A holistic consideration of authentic assessments: student perception of assessment design, delivery, flexibility and creativity. *European Journal of Engineering Education*, 51(1), pp. 25–42. <https://doi.org/10.1080/03043797.2025.2480116>
- Rawlasyk, P. (2018). Evaluación en la educación superior y el aprendizaje estudiantil. *Journal of Instructional Pedagogies*, Vol. 21. <https://www.aabri.com/manuscripts/182875.pdf>
- Rezai, A., Namaziandost, E., Miri, M. y Kumar, T. (2022). Demographic biases and assessment fairness in classroom: insights from Iranian university teachers. *Lang Test Asia* 12, 8. <https://doi.org/10.1186/s40468-022-00157-6>
- Richardson, M., Abraham, C. y Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138(2), pp. 353–387. <https://doi.org/10.1037/a0026838>
- Rutherford, S., Pritchard, C. y Francis, N. (2025), La evaluación ES aprendizaje: desarrollo de un enfoque centrado en el estudiante para la evaluación en la educación superior. *Biografía abierta de FEBS*, 15: pp. 21-34. <https://doi.org/10.1002/2211-5463.13921>
- Stankov, P. (2024). Are exams authentic assessment? The case of economics. *Practical Assessment, Research & Evaluation*, 29(3). <https://openpublishing.library.umass.edu/pare/article/1984/galley/1946/view/>
- Schoenfeld, A. H. (2016). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. *Journal of Education*, 196(2), pp. 1-38. <https://doi.org/10.1177/002205741619600202>
- Taylor, T. A. H., Kemp, K., Mi, M. y Lerchenfeldt, S. (2023). Prácticas de evaluación del aprendizaje autodirigido en la formación universitaria de profesiones sanitarias: una revisión sistemática. *Medical Education Online*. <https://doi.org/10.1080/10872981.2023.2189553>
- Thompson, P. W., Carlson, M. P., Byerley, C. y Hatfield, N. (2021). Schemes for thinking with magnitudes: A hypothesis about foundational reasoning abilities in algebra. In A. Weinberg, C. Rasmussen, J. Rabin, M. Wawro, & S. Brown (Eds.), *Proceedings of the 18th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 1-17). Mathematical Association of America.
- Tristán López, A. y Pedraza Corpus, N. Y. (2017). La objetividad en las pruebas estandarizadas. *Revista Iberoamericana de Evaluación Educativa*, 10(1), pp. 11–31. <https://doi.org/10.15366/riee2017.10.1.001>
- Vallejo Ruiz, M. y Molina Saorín, J. (2014). La evaluación auténtica de los procesos educativos. *Revista Iberoamericana de Educación*, 64, pp. 11–25.

<https://dialnet.unirioja.es/servlet/articulo?codigo=4774107>

Villarroel, V. y Bruna, D. (2019). ¿Evaluamos lo que realmente importa? El desafío de la evaluación auténtica en educación superior. *Calidad en la Educación*, 50, pp. 492–509.

<https://scispace.com/pdf/evaluamos-lo-que-realmente-importa-el-desafio-de-la-367g6p85x7.pdf>

Zueck, C., Rodríguez-Gil, A., Blanco-Vega, H., González-Rivas, R. A., & Núñez-Enríquez, O. (2023). Evaluación auténtica de aprendizajes: Acercamiento a una experiencia en universitarios de educación física. *Revista de Ciencias del Ejercicio FOD*, 18(2), pp. 35–44.

https://www.researchgate.net/publication/376996661_Evaluacion_autentica_de_aprendizajes_acercamiento_a_una_experiencia_en_universitarios_de_educacion_fisica