

Inteligencia Artificial: Implicaciones para la Evaluación en Educación Superior

Artificial Intelligence: Implications for Assessment in Higher Education

Pedro Antonio
Sánchez Escobedo¹

Diana
Otero Solís²

Resumen

Este artículo examina cómo la inteligencia artificial (IA) impacta la evaluación en la educación superior. Para ello, se analiza si las estrategias de evaluación, como exámenes orales y demostraciones, son útiles para fortalecer el pensamiento crítico y la capacidad de análisis frente a las herramientas digitales. Se presentan criterios para su uso y se destaca que la IA debe apoyar el aprendizaje, incentivando la reflexión, la toma de decisiones y la ejecución competente del alumnado.

Palabras clave: inteligencia artificial, educación superior, evaluación educativa.

Abstract

This article reflects on the implications of using Artificial Intelligence (AI) for evaluation in higher education. It is argued that qualitative evaluation strategies, such as oral exams and demonstrations, are valuable tools in fostering critical thinking, students' analytical abilities, and demonstrating competencies. Criteria for employing these evaluation techniques are reviewed. It is argued that while AI should support learning, it is essential that instructional procedures evidence the student's decision-making and performance.

Key words: artificial intelligence, higher education, educational assessment.

¹ Universidad Autónoma de Yucatán. Profesor Jubilado. Doctor. Yucatán, México. E-mail: psanchez@correo.uady.mx ORCID: <http://orcid.org/0000-0002-0564-3502> Google Scholar: <https://scholar.google.com/citations?user=yEpmNZwAAAAJ&hl=es>

² Universidad Autónoma de Yucatán. Asistente de investigación. Yucatán, México. E-mail: dianaoterosolis95@gmail.com ORCID: <https://orcid.org/0000-0002-4731-6366> Google Scholar: <https://scholar.google.com/citations?user=dF6c1ucAAAAJ&hl=es>



Introducción³

La evaluación constituye uno de los procesos fundamentales dentro de la educación superior, debido a que permite valorar el nivel de aprendizaje alcanzado por los estudiantes, así como la efectividad de los procesos de enseñanza. Tradicionalmente, la evaluación ha sido entendida como un mecanismo de medición del rendimiento académico mediante exámenes, trabajos escritos y actividades prácticas. Sin embargo, las tendencias educativas contemporáneas han impulsado una visión más integral, centrada en el desarrollo de competencias, el pensamiento crítico y la resolución de problemas (García y García, 2022).

Evaluar ante la implementación de la inteligencia artificial (IA) en educación superior es un reto vigente. La incorporación de la IA a la instrucción universitaria presenta desafíos abordados tanto a nivel nacional como internacional. Por ejemplo, la UNAM desarrollará el EvaluaFEST 2026 sobre “Evaluación del y para el aprendizaje con inteligencia artificial”, tratando temas como ética, pensamiento crítico y evaluación mediante IA (Consejo de Evaluación Educativa de la UNAM, 2026). A nivel global, instituciones como el Massachusetts Institute of Technology (MIT), la Universidad de Stanford y la Open University del Reino Unido, han comenzado a transformar sus sistemas de evaluación del aprendizaje mediante el uso de IA, incorporando herramientas que les permiten analizar el desempeño académico de una manera más personalizada, continua y eficiente (Hong et al., 2024; The Open University, 2026).

La IA es un conjunto de tecnologías que permiten que las computadoras realicen una

variedad de funciones avanzadas, por ejemplo, comprender y traducir lenguaje hablado y escrito, analizar datos a diferentes niveles, proponer soluciones, dar recomendaciones y mucho más. Se basa en la creación y aplicación de algoritmos para que los ordenadores “piensen” como los humanos (Nguyen, 2023).

Es un campo de la informática que se enfoca en crear sistemas que puedan realizar tareas complejas que emulan la inteligencia humana, como el razonamiento, la clasificación, el análisis de datos, el dibujo y diseño, la comprensión de idiomas, entre otras acciones antes reservadas exclusivamente a la mente humana (Bowman, 2022; García-Peñalvo, 2023). La inserción de la IA en todas las esferas de la vida humana y, especialmente, en la enseñanza es ineludible, por lo que resulta oportuno analizar sus implicaciones para la instrucción y particularmente para los procesos evaluativos.

La evaluación en educación

La evaluación consiste en un proceso sistemático destinado a recopilar información sobre el estudiante, con el propósito de verificar el cumplimiento de los objetivos establecidos en la instrucción. En otras palabras, es cualquier método utilizado para medir características de estudiantes, maestros, programas, cursos, etcétera. Esta representa una oportunidad para retroalimentar a todos los actores que forman parte del proceso educativo y debe constituir en sí misma una actividad de aprendizaje. La manera más efectiva de transformar el aprendizaje del estudiantado es modificar los métodos de evaluación (Elton

³ Este manuscrito fue revisado con ayuda de la inteligencia artificial (IA), específicamente el programa CopilotTM. Al comparar la versión original con la revisada se notaron cambios en la estructura y redacción, haciendo más sucinto e intencionado el documento. Se corrigieron dos párrafos en donde la idea original no fue respetada. El programa no eliminó o adicionó ninguna idea que no fuera de la propia autoría de los autores. También se extrajo información de ChatGPT respecto los métodos de evaluación.

y Laurillard, 1979). En la tradición ortodoxa, se realizan seis preguntas clave sobre evaluación:

Existen diversos tipos de evaluación, entre los cuales se encuentran la nomotética o de norma, que compara a los estudiantes entre sí para evidenciar su rendimiento relativo, y la ideográfica o de competencia, que valora al estudiante con base en un estándar preestablecido, destacando sus capacidades. Según el enfoque metodológico y

los resultados, la evaluación puede ser cuantitativa, utilizando datos numéricos y estadísticos para medir el desempeño grupal, o cualitativa, empleando descripciones y juicios para valorar el desempeño individual o colectivo (Espinoza, 2022).

En el contexto de la evaluación, se requiere que el lector recuerde con precisión los términos más comúnmente utilizados. La Tabla 2 presenta una recopilación de algunos de estos términos.

Tabla 1

Preguntas clave sobre la evaluación

Pregunta	Respuesta
¿Quién la hace?	Estudiantes (autoevaluación y coevaluación), docentes y autoridades educativas (heteroevaluación).
¿Qué se evalúa?	Los aprendizajes, habilidades, competencias y desempeño de los estudiantes, así como la calidad del proceso educativo.
¿Con que propósito?	Para monitorear el progreso, mejorar el aprendizaje y tomar decisiones educativas.
¿Dónde se hace?	En el aula, centro escolar y otros escenarios educativos presenciales o virtuales.
¿Cuándo se hace?	Puede realizarse previamente al desarrollo de un proceso educativo (diagnóstica), de manera sistemática (formativa) o al finalizar un proceso de instrucción o ciclo escolar (sumativa).
¿Cómo se hace?	Existe una diversidad de herramientas e instrumentos de evaluación como exámenes (orales y escritos), observaciones, proyectos, simulaciones, etcétera.

Fuente: elaboración propia.

Tabla 2

Términos esenciales en evaluación

Término	Definición
Medición	Consiste en asignar números o letras a una ejecución. Es parte del proceso de evaluación, una de las formas de hacer una evaluación cuantitativa.
Error estándar de medición (ESM)	Es la diferencia que existe entre el puntaje obtenido de la ejecución de un alumno y su “desempeño real” (el verdadero y teóricamente libre de contaminantes propios a toda observación). Se debe intentar reducir el ESM al mínimo, con la finalidad de acercarnos más al desempeño real del alumno.
Estimación	Es una apreciación subjetiva del desempeño del alumno, por lo general se hace en términos semánticos. Es instrumental para asignar una calificación. Es una de las formas de hacer una evaluación cualitativa.
Calificación	Es una letra, número o enunciado que es indicativo de la ejecución del alumno. Sirve para establecer rangos y jerarquías; para establecer si el alumno aprobó o no; y da una idea del grado de excelencia en el desempeño.
Prueba o examen	Es una muestra de la constelación de conductas o tareas esperadas en los alumnos y descritas en los objetivos instruccionales. Las pruebas tienen por objeto estimar el grado en que el alumno han alcanzado los objetivos. Pueden ser escritas o rutinas de ejecución psicomotora (por ejemplo, la prueba de manejo). No son los únicos recursos para evaluar, muchas tareas, trabajos, participaciones y observaciones pueden servir para hacer una evaluación.

Fuente: elaboración propia.

Desarrollo

¿Qué implicaciones concretas tiene la IA para la evaluación?

La IA presenta desafíos significativos para el docente, quien deberá afrontar la incorporación intencionada de esta tecnología para optimizar la instrucción. Lo anterior exige de capacitación digital, pero también una planificación cuidadosa de su utilización y un enfoque equilibrado entre los procesos digitales y mentales. En este sentido, reflexionar sobre las implicaciones de la IA en la evaluación es imperativo para potenciarla como un recurso que favorezca el aprendizaje (Incio et al., 2022; Qu, Zhao y Xie, 2022).

La aparición de herramientas de IA generativa ha cuestionado los métodos tradicionales de evaluación utilizados en las universidades. Actividades

como ensayos, reportes escritos o cuestionarios realizados fuera del aula pueden ser elaborados parcial o totalmente mediante IA, dificultando la identificación del nivel real de aprendizaje del estudiante. Ante esta situación, las estrategias evaluativas deben evolucionar hacia modelos más auténticos y centrados en la verificación de las competencias del estudiante.

La evaluación, en el marco de la IA, deberá priorizar la aplicación del conocimiento en contextos reales, promoviendo actividades donde el estudiante demuestre objetivamente su capacidad de análisis, creatividad, argumentación y de resolución de problemas. En este contexto, cobran relevancia estrategias como las evaluaciones orales, los estudios de caso, proyectos integradores, aprendizaje basado en problemas, portafolios digitales, actividades colaborativas y evaluaciones prácticas supervisadas.

Tabla 3

Tipos de exámenes orales

Tipo	Características
Tradicional	Consiste en una serie de preguntas formuladas por el docente de manera individual al estudiante. Generalmente se realiza frente a un jurado o un profesor evaluador y busca comprobar conocimientos teóricos, comprensión conceptual y capacidad de análisis.
Exposición oral	El estudiante presenta un tema previamente asignado o investigado. Este método evalúa tanto el dominio del contenido como la organización de ideas, la argumentación y la expresión verbal.
Defensa de proyectos o tesis	Es común en licenciatura y posgrado. El estudiante expone y defiende un trabajo académico ante un comité evaluador, respondiendo preguntas relacionadas con metodología, resultados y conclusiones.
Examen oral estructurado	Se basa en preguntas previamente definidas y criterios de evaluación estandarizados. Su propósito es reducir la subjetividad y garantizar mayor equidad entre los estudiantes.
Entrevista académica	Se desarrolla en formato de diálogo entre evaluador y estudiante. Permite explorar procesos de razonamiento, reflexión y aplicación práctica del conocimiento.
Evaluación mediante estudio de casos	El estudiante analiza y responde oralmente a situaciones problemáticas o escenarios reales. Este método es frecuente en áreas como medicina, derecho, administración y educación.
Exámenes orales grupales	Se realizan con varios estudiantes simultáneamente y suelen emplearse para valorar habilidades colaborativas, discusión académica y argumentación colectiva.

Fuente: elaboración propia.

A continuación, se analizan dos técnicas de uso frecuente: evaluaciones orales y demostraciones en vivo. Estas modalidades reducen la influencia de la IA en la comprobación del aprendizaje y la ejecución del estudiante.

Exámenes orales: tipos y normativas

Los exámenes orales permiten valorar de manera directa el dominio conceptual, la capacidad argumentativa, el pensamiento crítico y las habilidades comunicativas del estudiante; se vislumbran como una alternativa eficaz frente al plagio y al uso inadecuado de herramientas de IA, debido a que permiten verificar en tiempo real la comprensión auténtica del estudiante. A diferencia de los exámenes escritos, este tipo de evaluación favorece la interacción inmediata entre docente y alumno, posibilitando la profundización en las respuestas y la aclaración de ideas. La Tabla 3 resume diversos tipos de exámenes orales y sus características.

Para garantizar objetividad, transparencia y validez en los exámenes orales, las instituciones educativas suelen establecer lineamientos normativos y pedagógicos. Es indispensable que los estudiantes conozcan previamente los objetivos del examen, los contenidos a evaluar, la duración, la modalidad y, sobre todo, los criterios de calificación. Frecuentemente se utilizan rúbricas o escalas de desempeño para disminuir la subjetividad. Todos los estudiantes deben ser evaluados bajo condiciones similares, por ejemplo, mismos tiempos, criterios, acceso y preguntas con dificultad equivalente. El evaluador debe mantener una actitud imparcial, respetuosa y profesional, evitando preguntas ambiguas, intimidatorias o fuera de los contenidos establecidos.

Ante el uso cada vez más extendido de la IA en la educación, los exámenes orales cobran relevancia como forma de evaluar la comprensión, el pensamiento crítico y la autoría del aprendizaje. Por ello, muchas universidades suman evaluaciones orales a las actividades escritas realizadas con tecnología. No obstante, esta implementación exige docentes capacitados, criterios claros y mecanismos para reducir la subjetividad y garantizar procesos justos. Muchas instituciones sugieren registrar el examen con actas, grabaciones o notas del jurado, esto permite transparencia y posibilidad de revisión en caso de inconformidades. Si hay un solo evaluador, debe justificar su calificación por escrito y de forma cualitativa.

Cuando el jurado está compuesto por dos o más docentes, la calificación se otorga por consenso grupal, sin considerar criterios individuales. Así, el estudiante recibe retroalimentación colectiva, no de una persona específica. Además del conocimiento disciplinar, suelen evaluarse aspectos como claridad de expresión, coherencia argumentativa, uso adecuado del lenguaje técnico y capacidad de síntesis. Es fundamental que los evaluadores no premien la emotividad o la vehemencia por encima de los argumentos. Se debe evitar valorar en exceso la retórica, entendida como la teoría para estructurar discursos persuasivos, y la oratoria, que es su aplicación práctica para hablar con elocuencia y convencer al público. Mediante el discurso se evalúan la estructura, veracidad, originalidad, coherencia y profundidad del pensamiento del estudiante.

Demostraciones como método de evaluación

Las demostraciones constituyen un método de evaluación centrado en la ejecución práctica de conocimientos, habilidades y competencias por

parte del estudiante. A través de este tipo de evaluación, el alumno evidencia de manera directa su capacidad para aplicar procedimientos, resolver problemas o desarrollar tareas específicas en contextos reales o simulados.

Este método es ampliamente utilizado en educación superior, especialmente en áreas donde el aprendizaje práctico resulta fundamental, como ciencias de la salud, ingeniería, educación, artes, tecnología y formación técnica. Las demostraciones permiten valorar no solo la teoría, sino también la capacidad de ejecución. En estas, el estudiante muestra cómo realiza una actividad o procedimiento, es decir aplica los conocimientos en un contexto real o simulado y permite que el evaluador observe, registre, y retroalimente su ejecución. Existen cinco tipos comunes de demostraciones como mecanismos de evaluación en la educación superior (ver Tabla 4).

Las demostraciones permiten evaluar competencias reales y no solamente la teoría (memoria); favorecen el aprendizaje aplicado y se asocian estrechamente a las competencias profesionales requeridas en el mundo laboral;

facilitan la evaluación de competencias complejas, como comunicación, liderazgo y resolución de problemas, reduciendo las posibilidades de plagio o dependencia excesiva de la inteligencia artificial.

A pesar de sus ventajas, las demostraciones presentan algunos desafíos, por ejemplo, requieren mayor tiempo de evaluación, además de que pueden necesitarse recursos materiales y escenarios especializados, y existe el riesgo de subjetividad en la observación del evaluador. Por otra parte, la ejecución siempre está contaminada por aspectos extracurriculares como la personalidad, el grado de ansiedad y la competencia social del estudiante. Por ello, se recomienda utilizar instrumentos claros como rúbricas, listas de cotejo o escalas de desempeño, en los que se deben clarificar los aspectos a medir, así como el grado de dominio y precisión esperados.

Las demostraciones se consideran como una forma de evaluación auténtica, ya que permiten valorar el desempeño del estudiante en tareas similares a las que enfrentará en el ámbito profesional. Por ello, diversas instituciones de educación superior promueven el uso de

Tabla 4

Tipos de demostraciones

Tipo	Descripción
1. Demostración práctica	El estudiante ejecuta procedimientos o técnicas específicas frente al evaluador, por ejemplo, realizar un experimento, utilizar herramientas tecnológicas o aplicar estrategias didácticas.
2. Demostración experimental	Consiste en comprobar principios científicos o procesos mediante experimentos o prácticas de laboratorio.
3. Simulación	El estudiante actúa en escenarios similares a situaciones reales, como entrevistas, atención a pacientes, juicios orales o resolución de conflictos.
4. Presentación de proyectos	Incluye la exposición y demostración funcional de productos, prototipos, investigaciones o propuestas desarrolladas por el estudiante.
5. Microenseñanza	Utilizada en formación docente, donde el estudiante demuestra habilidades pedagógicas mediante clases simuladas.

Fuente: elaboración propia.

demostraciones, simulaciones y evaluaciones prácticas como estrategias complementarias a los métodos tradicionales de evaluación.

Discusión

Los profesores cuentan con el gran recurso de la evaluación para mitigar los riesgos de la IA para el aprendizaje. Los métodos tradicionales de evaluación, como los exámenes orales, las pruebas de selección múltiple, los ensayos manuscritos y la ejecución libre de aditamentos digitales, deberán de tomar relevancia en la vida académica para garantizar que los estudiantes piensan, juzgan y toman decisiones sin depender de las plataformas que utilizan.

Es necesario diseñar estrategias de evaluación en vivo, que evidencien, sin la influencia de la IA, el aprendizaje adquirido y el desarrollo de las competencias preestablecidas en el currículo. Los exámenes orales y las demostraciones, como se mencionó, permiten fomentar el pensamiento crítico al complementarse con la amplia información y las capacidades de análisis que pueden proporcionar las herramientas de IA. La evaluación presencial favorece la integridad académica al incentivar actividades auténticas.

De igual manera, es importante considerar que la integración de la IA en la labor de enseñanza puede facilitar el trabajo docente y abrir nuevas posibilidades para el diseño de las actividades evaluativas. Al incorporar estas herramientas, los docentes pueden ofrecer evaluaciones más ricas y personalizadas, que no solo miden el conocimiento, también fomentan el desarrollo integral de los estudiantes (Bauer, et al., 2025; Ramón, 2021). La evaluación en el ámbito educativo, a través

de la IA, ofrece a los docentes una variedad de herramientas para mejorar la precisión, la equidad y la eficacia de la medición del aprendizaje logrado; particularmente, cuando utilizan métodos cualitativos de evaluación con criterios adecuados y parámetros objetivos y claros.

No se trata de prohibir el uso de la IA, sino de impulsar una alfabetización digital crítica, enseñando a emplearla de manera ética y responsable. Para ello, es imprescindible que las instituciones establezcan reglas claras para su uso en contextos educativos (Chan, 2023).

Conclusiones

La IA exige que la evaluación en educación superior se enfoque en desarrollar habilidades como pensamiento crítico, creatividad y análisis, más allá de memorizar contenidos. El docente debe generar experiencias de evaluación vinculadas a situaciones reales para evidenciar el aprendizaje del alumnado. En este sentido, la evaluación debe reconocer que la IA formará parte del entorno profesional de los estudiantes y evaluar más allá de lo aprendido (evaluación sumativa), sino también la capacidad del estudiante para utilizar estas herramientas de manera reflexiva, ética y estratégica (evaluación formativa).

La educación superior enfrenta así el desafío de construir modelos evaluativos más flexibles, humanistas y centrados en competencias, capaces de responder a las transformaciones derivadas del avance tecnológico, teniendo como reto decantar el apoyo de la tecnología y evidenciar lo que existe en la mente (ideas) y el cuerpo (habilidades) de los estudiantes.

Referencias

- Bauer, E., Greiff, S., Graesser, A. C., Scheiter, K. y Sailer, M. (2025). Looking Beyond the Hype: Understanding the Effects of AI on Learning. *Educational Psychology Review*, 37(45), 1-27. DOI: <https://doi.org/10.1007/s10648-025-10020-8>
- Bowman, E. (2022). *A new AI chatbot might do your homework for you. But it's still not an A+ student*. NPR. <https://www.npr.org/2022/12/19/1143912956/chatgpt-ai-chatbot-homework-academia>
- Chan, C. K. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(38), 1-25. DOI: <https://doi.org/10.1186/s41239-023-00408-3>
- Consejo de Evaluación Educativa de la UNAM (2026, 16 de abril). EVALUAFEST 2026. *Gaceta UNAM*, p. 1.
- Elton, L. R. y Laurillard, D. M. (1979). Trends in Research on Student Learning. *Studies in Higher Education*, 4(1), 87-102. DOI: <http://dx.doi.org/10.1080/03075077912331377131>
- Espinoza Freire, E. E. (2022). La evaluación de los aprendizajes. *Conrado*, 18(85), 120-127. http://scielo.sld.cu/scielo.php?pid=S1990-86442022000200120&script=sci_arttext&tlng=pt
- García Acosta, J. G. y García González, M. (2022). La evaluación por competencias en el proceso de formación. *Revista Cubana de Educación Superior*, 41(2), 1-19. http://scielo.sld.cu/scielo.php?pid=S0257-43142022000200022&script=sci_arttext
- García-Peñalvo, F. J. (2023). La percepción de la Inteligencia Artificial en contextos educativos tras el lanzamiento de ChatGPT: disrupción o pánico. *Education in the Knowledge Society (EKS)*, 24, e31279. <https://doi.org/10.14201/eks.31279>
- Hong, S., Cai, C., Du, S., Feng, H., Liu, S. y Fan, X. (2024). “MyGrade is Wrong!”: A Contestable AI Framework for Interactive Feedback in Evaluating Student Essays. *ArXiv*, 1-9. DOI: <https://doi.org/10.48550/arXiv.2409.07453>
- Incio Flores, F. A., Capuñay Sanchez, D. L., Estela Urbina, R. O., Valles Coral, M. A., Vergara Medrano, S. E. y Elera Gonzales, D. G. (2022). Inteligencia artificial en educación: una revisión de la literatura en revistas científicas internacionales. *Apuntes Universitarios*, 12(1), 353-372. DOI: <https://doi.org/10.17162/au.v12i1.974>
- Nguyen, N. D. (2023). Exploring the role of AI in education. *London Journal of Social Sciences*, 6(6), 84-95. DOI: <https://doi.org/10.31039/ljss.2023.6.108>
- Qu, J., Zhao, Y. y Xie, Y. (2022). Artificial intelligence leads the reform of education models. *Systems Research and Behavioral Science*, 39(3), 581-588. DOI: <https://doi.org/10.1002/sres.2864>
- Ramón Pucurucu, L. P. (2021). Beneficios y desventajas del uso de la Inteligencia Artificial (IA) en educación superior. *Revista Social Fronteriza*, 1-13. DOI: [https://doi.org/10.59814/resofro.2021.1\(1\)252](https://doi.org/10.59814/resofro.2021.1(1)252)
- The Open University (2026). *Learning analytics*. The Open University. <https://help.open.ac.uk/index.php/learning-analytics>