

Experior: Revista de Investigación de ADEN University
ISSN L 2953-3090
Vol. 3 (1) enero-junio 2024

Aprender y enseñar en la universidad: AI Generativa en los procesos de evaluación

Learning and teaching at the university: Generative AI in assessment processes

Dairaliz Marcano de Leal
Pontificia Universidad Católica Argentina, Argentina
dairalizmarcano@gmail.com
<https://orcid.org/0000-0003-2216-9010>

Recibido: 15/03/2024.

Aceptado: 15/05/2024.

Publicado: 30/06/2024.

Cómo citar: Marcano de Leal, D. (2024). Aprender y enseñar en la universidad: AI Generativa en los procesos de evaluación. *Experior*, 3(1), 59-69.
<https://doi.org/10.56880/experior31.6>

Resumen

El estudio tuvo por objetivo analizar los cambios y oportunidades que plantea la IA generativa a los procesos de evaluación en el marco de la enseñanza y el aprendizaje en la universidad. Bajo el paradigma de investigación cualitativa, se realizó una investigación documental. Se recopilaron estudios que analizan la irrupción de la IA en los procesos de evaluación de los aprendizajes en la educación universitaria. Además, se analizaron documentos elaborados por doctorandos utilizando una herramienta de IA generativa. Desde los fundamentos conceptuales que explican el aprendizaje profundo y las nuevas pedagogías activas, el uso de la IA supone resignificar los propósitos, fines y objetivos de la evaluación de los aprendizajes. Los resultados confirman la necesidad de incorporar el uso de las IA alineadas con objetivos de aprendizaje complejos, contextualizados y auténticos, la definición de criterios de transparencia de intencionalidad pedagógica para estudiantes y docentes y garantías para el uso adecuado y acorde a principios éticos. Se proponen diversas alternativas para que las universidades incorporen las IA, como el establecimiento de lineamientos institucionales y elaboración de escalas de evaluación con criterios para la inclusión y uso de la IA generativa.

Palabras clave: IA generativa, Aprendizaje profundo, Nuevas pedagogías, Educación universitaria, Evaluación.

Abstract

The study was aimed to analyze the changes and opportunities that generative AI poses to assessment processes in the framework of teaching and learning at the university. Under the qualitative research paradigm, documentary research was conducted. Studies were compiled that analyze the emergence of AI in the learning assessment processes in university education. In addition, documents prepared by doctoral students using a generative AI tool were analyzed. From the conceptual foundations that explain deep learning and new active pedagogies, the use of AI involves redefining the purposes, aims and objectives of learning assessment. The results confirm the need to incorporate the use of AI aligned with complex, contextualized and authentic learning objectives, the definition of criteria for transparency of pedagogical intentionality for students and

teachers and guarantees for proper use and in accordance with ethical principles. Various alternatives are proposed for universities to incorporate AI, such as establishing institutional guidelines and developing evaluation scales with criteria for the inclusion and use of generative AI.

Keywords: Generative AI, Deep learning, New pedagogies, University education, Evaluation.

Introducción

La naciente era digital cambia aspectos fundamentales de la educación. Cambia los roles tradicionales de los profesores y los libros de texto como fuentes primarias de conocimiento del contenido. En el pasado, para los educadores, el término "aplicar conocimientos" significaba trabajar en tareas o resolver problemas para demostrar dominio de conceptos. Los docentes asumen nuevos roles como mediadores, mentores y creadores de ambientes de aprendizaje.

Cambia lo que es posible para los estudiantes, ya que la tecnología posibilita descubrir, crear y utilizar conocimientos en el mundo real más rápido, más inmediato y postular sus soluciones a problemas del mundo real con audiencias auténticas bien más allá de los límites de sus escuelas (Caballero, 2022). Este es el potencial real de la tecnología para afectar el aprendizaje, no para facilitar la entrega y consumo de conocimiento, sino para permitir a los estudiantes utilizar, crear, cuestionar y validar su conocimiento en el mundo (Fullan & Langworthy, 2014).

Con la incorporación de la inteligencia artificial (IA) en el ámbito educativo, se suman nuevos desafíos a los ya planteados por la educación digital, con renovadas oportunidades para transformar las prácticas tradicionales de enseñanza y evaluación. Todo indica que el avance de estas tecnologías acerca la posibilidad de personalizar los aprendizajes, automatizar procesos y ofrecer retroalimentación inmediata a los estudiantes (Furze, 2024). Sin embargo, este creciente protagonismo de la IA en la educación también plantea interrogantes cruciales sobre sus limitaciones y su impacto en la evaluación de los aprendizajes. ¿Puede la IA captar la complejidad y riqueza del proceso de aprendizaje o corre el riesgo de reducir la evaluación a un conjunto de datos medibles?

Con base a la revisión documental de investigaciones que analizan el impacto de las IA generativa en los procesos de evaluación, el artículo explora tanto las capacidades como los desafíos, a la luz de los avances en la comprensión de los procesos de aprendizaje y su conexión con nuevas metodologías de enseñanza. Se propone una revisión crítica del papel de la tecnología en la evaluación de los resultados de aprendizaje, y se describen alternativas que favorecen la inclusión de las IA generativa con intencionalidad pedagógica orientada al aprendizaje profundo, alineada a resultados de aprendizaje, criterios transparentes y conciencia de los límites que imponen frente a los amplios y diversos espacios de acción donde la práctica docente resulta insustituible.

Método

Bajo el paradigma cualitativo, se propuso una investigación documental, con el objetivo de analizar los cambios y oportunidades que ofrece la IA generativa a los procesos de evaluación en el contexto de la educación universitaria.

El diseño incluyó la recolección de publicaciones académicas a partir del 2022 al 2024. Se contemplaron como criterios de inclusión investigaciones o artículos de revisión de revisión teórica con los descriptores IA generativa, educación universitaria y evaluación de los aprendizajes. Se utilizaron las herramientas EBSCON y SCISPACE, y se consideró la utilización de la red social LinkedIn dado que los investigadores publican reflexiones que resultaban de interés a los fines de la investigación. En total se

seleccionaron cinco artículos. También se analizaron treinta y dos respuestas a un foro de discusión virtual de participantes de un curso de doctorado en Educación Virtual. Se utilizó la técnica de análisis de contenido.

Resultados

Aprendizaje profundo y nuevas pedagogías

La investigación del aprendizaje del estudiante ha descrito dos formas de aprender una determinada tarea (Maton & Säljö, 1976 como fue citado por Bigg, 2006) denominados enfoque superficial y profundo. En el enfoque superficial, la intencionalidad del estudiante es liberarse de la tarea con el mínimo esfuerzo, aunque cuidando impresionar con que se satisfacen los requerimientos académicos impuestos, se concentra en la adquisición de una lista de datos/conceptos sin comprender su significado y estructura, tratando de recordar tanta información como sea posible y reproducirla de manera admisible, poniendo en acción procesos cognitivos de nivel básico.

Las nuevas pedagogías pueden definirse sucintamente como un nuevo modelo de asociaciones de aprendizaje entre estudiantes y profesores, con el objetivo de alcanzar objetivos de aprendizaje profundo y habilitados por acceso digital. En el enfoque profundo, el estudiante siente la necesidad de utilizar los procesos cognitivos apropiados para desarrollar la tarea de forma comprensiva y significativa, centrado en descubrir el significado subyacente relacionado con los aprendizajes previos de manera de distinguir entre las evidencias y las conclusiones. El aprendizaje despierta en el estudiante interés, sentido de importancia, sensación de desafío; aprender es un placer y se muestran motivados a hacer preguntas y buscar respuestas utilizando procesos cognitivos de orden superior.

Si para la ciencia cognitiva el estudiante es capaz de experimentar interés y motivación para aprender, ¿por qué la primera respuesta de los docentes y sistemas educativos frente a la irrupción de la IA generativa es de una profunda desconfianza? Palabras como “trampa” y “plagio”, presentes en el ideario sobre lo que es y no evaluar en el ámbito de la educación, se han visto avivadas por las potencialidades que ofrecen las IA generativa para resolver las tareas que tradicionalmente utilizamos para valorar el logro de los objetivos de aprendizaje.

Más allá de la discusión de si el uso de IA generativa se considera plagio o trampa, complicada por las inconsistencias observadas en las herramientas para detectar productos aplicaciones como CHAT GPT 4, lo que demuestran las investigaciones es que cada vez más estudiantes hacen uso, apropiado o no, de estas herramientas para completar las tareas en las que se espera se evidencien los aprendizajes alcanzados (Perkins et al, 2024).

Evaluar en la universidad en tiempos de IA generativa

Lo cierto es que existen suficientes evidencias de la capacidad de las herramientas IA basadas en texto para generar artefactos indetectables por los docentes, quienes pueden aprobarlos sin que represente un aprendizaje real de los estudiantes.

Ante la IA generativa, la respuesta de los sistemas educativos ha sido diversa. Lodge et al (2023) propone una taxonomía que parte de ignorar su presencia en las aulas y esperar que, como otras tecnologías, desaparezca su interés por parte de los estudiantes. La evidencia demuestra que esta opción es cada vez más inviable a mediano y largo plazo. Todo indica que la IA generativa tendrá cada vez un impacto significativo en la educación.

Otros sistemas han optado por prohibir su uso. Nuevamente, esta opción demostró muy pronto su escasa viabilidad ante lo accesible que resultan los métodos

para burlar de manera eficaz los detectores de IA. Además, esta opción coloca a los docentes en la penosa tarea de dedicar esfuerzos infructuosos para detectar trabajos elaborados con herramientas y refuerza la imagen de los estudiantes como deshonestos y tramposos, que, si bien los hay, afortunadamente no son la mayoría. Tal como se afirma en párrafos precedentes, con tareas debidamente desafiantes y orientada con las necesidades e interés, los estudiantes están dispuestos a realizar las acciones cognitivas que conllevan a aprender con profundidad.

Hay quienes abogan por métodos de vigilancia como regresar a las pruebas de lápiz y papel, pruebas presenciales y otras estrategias donde los estudiantes produzcan evidencias escritas y orales de lo aprendido. Si bien estos métodos pueden generar mayor confianza en que reflejan el aprendizaje, son apropiados en algunas disciplinas y contenidos, requieren estar bien diseñados y contar con espacios adecuados para su efectividad.

Otros sistemas optaron por incluir el uso de la IA generativa en los procesos de enseñanza en general, y en las evaluaciones de manera particular. Diversas iniciativas se proponen en este sentido. Un ejemplo de esta mirada se observa en la *Guía sobre el Uso de la Inteligencias Artificial en el ámbito educativo (2024)*, del Gobierno de España, que identifica tres ámbitos para integrar su uso en los procesos de aprendizaje de: a) habilidades y competencias relevantes para los desafíos y oportunidades que la IA (Enseñar para las IA); b) conocimientos técnicos para aplicar y desarrollar herramientas IA (Enseñar sobre la IA) y c) personalización y gestión del entorno educativo (Enseñar con la IA). Experiencias como realizar tareas específicas de construcción de instrucciones, criticar productos, evaluar la veracidad de los artefactos generados, son algunos ejemplos de la adopción de las IA.

Con la inclusión de la IA, resulta insoslayable debatir los aspectos éticos, la justicia y equidad en las oportunidades de uso, en términos de establecer mecanismos que garanticen el acceso de tecnologías avanzadas y las capacidades para el uso eficiente por parte de estudiantes y docentes, frente al riesgo de que se profundicen las brechas digitales.

Desde la irrupción del Chat GPT en noviembre de 2022, quien escribe este artículo militó en la quinta opción: proponer actividades de evaluación que lleven a los estudiantes a pensar, en el entendido de que la IA generativa presenta limitaciones para el pensamiento crítico y reflexivo, para resolver problemas en contexto situados y en tareas auténticas. Si como afirma Ritchhart, et al (2014) aprender es pensar y pensar es aprender, entonces "*hacer visible el pensamiento* es de gran importancia para la educación" (p. 14) y es tarea fundamental del proceso de evaluación. Desde este paradigma, la cultura del pensamiento diferencia el *pensar acerca de* datos y conceptos, y *pensar con* los datos y conceptos para dotar de significado situaciones, problemas y hechos que están en la realidad situada del estudiante.

El aprendizaje profundo se centra en el desarrollo de la comprensión a través de procesos más activos y constructivos, resultado de acciones cognitivas como cuestionar, documentar, nombrar, entre otras. Implica poner el foco en el cuidadoso diseño de actividades situadas en su contexto disciplinar; comprometer al estudiantado en tareas que involucran la aplicación de conocimientos, habilidades y capacidades en situaciones complejas, realistas y contextualizadas.

"Esto significa resolver problemas, tomar decisiones y desarrollar nuevas comprensiones utilizando los métodos y las herramientas de la disciplina. Tenemos que ser conscientes de los tipos de pensamiento que son importantes para los científicos (hacer y probar hipótesis, observar de cerca, construir explicaciones...), matemáticos (buscar patrones, hacer conjeturas, formular generalizaciones, construir argumentos...), lectores (hacer interpretaciones, conexiones, predicciones...), historiadores (considerar diferentes puntos de

vista, razonar con evidencia, construir explicaciones...) y así sucesivamente.” (Ritchhart, et al, 2014, p 27).

En la Figura 1 se muestra una de las actividades diseñadas para un foro de discusión virtual en un curso de doctorado en Educación Virtual.

Figura 1.

Actividad foro de discusión de un curso doctoral en Educación Virtual.

Tema de discusión. Complemento humano a la tecnología.

Se propone la rutina de pensamiento **Tomar postura:**

Punto de partida:

Para leer: La Comisión Internacional Sobre Futuros de la Educación (2022) afirma que: “Ante la creciente omnipresencia de los sistemas automatizados que emplean IA y que prometen rutas ya diseñadas para la enseñanza, el aprendizaje o la evaluación, necesitamos un complemento humano y un contrapeso”.

Para ver: En este vídeo (enlace), Kai-Fu Lee analiza cómo las nuevas tecnologías pueden mejorar la labor educativa y el trabajo de los profesores.

Tomar postura:

Sustenta una posición convergente o divergente con respecto a lo visto y leído en los puntos de partida. ¿Qué te hace pensar esto? ¿Cuáles argumentos sustentan tu perspectiva? Incluye al menos dos participaciones argumentando tu respuesta frente a la postura de otros participantes. Fundamentar con citas y referencias bibliográficas.

En un ejercicio práctico, quien escribe solicitó a Chat GPT4 responder a la consigna del foro. La narrativa resulta en un texto que si bien fija una postura, la argumentación reproduce la idea de los dos puntos de entrada, sin ahondar en una reflexión que argumente la postura personal. Además, no puede incluir participaciones previas, lo que requiere una acción intencionada del estudiante para reconocer puntos de convergencia y divergencia con otros.

Figura 2

Instrucción y respuesta a la consigna en CHAT GPT4.

Actúa como un estudiante de doctorado en Educación Virtual. Resuelve la siguiente consigna en un texto fluido, en no más de 1000 caracteres. Lee la consigna en la captura de pantalla y ten en cuenta todos los aspectos condicionantes de la tarea.

Aquí tienes una versión mejorada de la redacción:

La afirmación de la Comisión Internacional Sobre Futuros de la Educación (2022) enfatiza la necesidad de un equilibrio entre la inteligencia artificial (IA) y la intervención humana en la educación. Estoy de acuerdo con esta perspectiva, pues, aunque la IA facilita procesos eficientes en la enseñanza y evaluación, es imperativo que el factor humano permanezca central en el ámbito educativo. Los docentes no solo aportan la empatía y el juicio ético necesarios para abordar las complejidades emocionales y sociales de los estudiantes, sino que también actúan como un contrapeso vital frente a los posibles sesgos y limitaciones de la tecnología.

Kai-Fu Lee subraya cómo las tecnologías emergentes pueden mejorar el trabajo docente, pero sostengo que la tecnología debe ser vista como una herramienta complementaria, no como un reemplazo. La clave está en lograr una integración armónica donde ambos elementos se potencien mutuamente, garantizando que la educación continúe siendo un proceso profundamente humanizador y no meramente técnico.

Nota: OpenAI. (2024).

En total, se obtuvieron treinta y dos respuestas al foro. La totalidad de las entradas fueron evaluadas a través del detector Quillbot (<https://quillbot.com/es/detector-de-ia>). Los resultados (Tabla 1) indican que solo una respuesta fue evaluada por el detector como 100% generado por humano mientras que un 56, 25% calificaron con un porcentaje de escritura con IA superior a los 51%. ¿Podemos confiar en los porcentajes arrojados por el detector? La evaluación realizada de las respuestas dio cuenta de lo errado de los porcentajes descritos. Un porcentaje significativamente menor evidenció falencias similares a las encontradas en la respuesta dada por Chat GPT 4, mientras que la mayoría mostró un desarrollo de ideas con base argumentativa sólida a favor de la postura sostenida, con elaboración propia de argumentos más allá de los descritos en las fuentes que sirvieron de punto de partida.

Tabla 1.

Porcentaje de texto generado por IA.

Generado por IA	Respuestas	
1 a 50%	14	43,75%
51 a 100%	17	56,25

Si bien esta opción resulta prometedora, es posible que, con los avances de la IA generativa, sea insuficiente proponer actividades de evaluación que valoren el producto, aun considerando lo expuesto antes: tareas centradas en demostrar la comprensión de los conceptos, habilidades y capacidades en situaciones concretas (Lodge et al 2023).

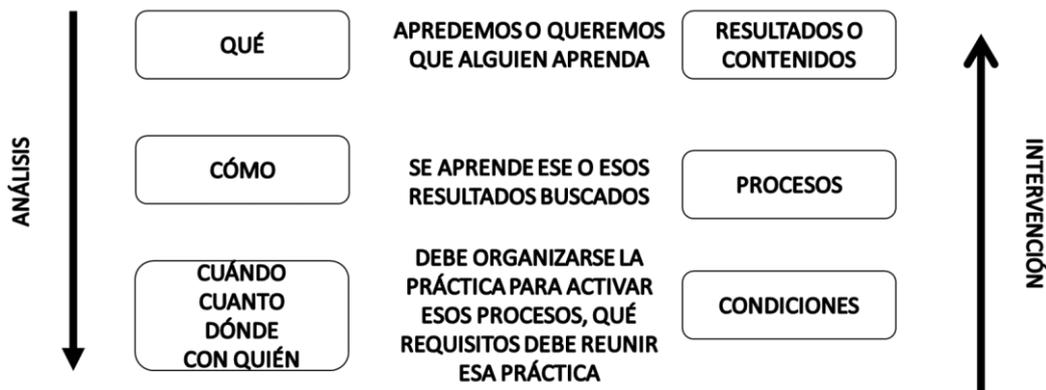
Repensar la evaluación

Si bien las actividades que buscan bordear los límites de la IA pueden abonar a una evaluación más auténtica de los aprendizajes, es posible que sigan proporcionando instantáneas discretas del desempeño en lugar de visiones matizadas del aprendizaje. Repensar la evaluación se presenta como la opción con mayor viabilidad a futuro.

Al ser el aprendizaje un proceso interno, invisible al ojo del docente, y en la mayoría de los casos, al mismo estudiante, se infiere a partir de los *resultados de aprendizaje*, que expresa lo que se espera que los estudiantes sepan, comprendan y sean capaces de hacer al término de un período de aprendizaje (contenidos – capacidades). Dentro del paradigma de la comprensión, los *procesos de aprendizaje*, referidos a la actividad mental de la persona que hace posible los cambios (cómo se aprende), se buscan apuntalar a partir de estrategias como las rutinas de pensamiento, guiando a los estudiantes a hacer visible cómo van desarrollando el aprendizaje. Es decir, a partir del arreglo cuidadoso de las *condiciones del aprendizaje*, se propone el tipo de práctica que pone en marcha dichos procesos de aprendizaje.

Aprender es un proceso constituido por diferentes momentos y fases, encadenadas en secuencias lógicas de las que depende el logro de los resultados de aprendizaje (Zabalza, 2022). Sin embargo, durante el proceso de instrucción o capacitación, lo que podemos observar y medir es el desempeño, que a menudo es un índice poco confiable de si se han producido los cambios a relativamente largo plazo que constituyen el aprendizaje (Soderstrom & Björk, 2015).

Figura 3.
Situaciones de aprendizaje.



Nota: Pozo, 2001.

Cada vez más voces abogan por repensar la evaluación que considere el aprendizaje como proceso complejo, personalizado y secuenciado. La enseñanza efectiva debe articular los dispositivos dispuestos por el docente con el proceso de aprendizaje y las fases que entraña en el recorrido que debe realizar el estudiante hasta llegar a un evidenciar un resultado de aprendizaje. Cómo conectar los dispositivos de enseñanza con el proceso de aprendizaje, continúa siendo un aspecto poco atendido en la docencia.

Resulta una tarea pendiente para los docentes profundizar en la comprensión del aprendizaje que esperamos logren nuestros estudiantes y cómo estructurar los artefactos mediacionales para acompañarlos por el tránsito que representa llegar a aprender. Responder a las preguntas qué aprendizajes son significativos, con qué se espera sean evidentes y cómo documentarlos para su evaluación, son preguntas que los docentes debemos incorporar como punto de partida al planificar los procesos de enseñanza y aprendizaje.

Tomando como base lo expuesto por Zabalza (2022), con base a una investigación de Oser y Baeriswyl (2001), y a modo de mostrar la complejidad de la tarea mencionada, se reproduce una clasificación de diversos formatos de aprendizaje, estructurada con base a lo que se espera que los estudiantes aprendan, y que matiza el cómo alcanzarlo y también el resultado de aprendizaje deseado.

Tabla 2.
Formatos de aprendizaje.

Categoría	Sentido del aprendizaje	Modelos base incluido
Aprendizaje cognitivo	Aprender es adquirir conocimiento y saber utilizarlos para comprender la realidad	Descubrimiento Solución de problemas Conocimiento semántico Aprendizaje de conceptos Aprender a aprender
Aprendizaje experiencial	Aprender es perfeccionar la capacidad de ejecutar conductas en contextos sociales.	Rutina y habilidades Aprendizaje de motilidad Aprendizaje social Aprendizaje Hipertexto

		Aprendizaje negociación
Aprendizaje valores actitudes	y	Desarrollo moral
Aprender es adquirir nuevas formas de ver la vida y afrontar las relaciones		Reflexión y contemplación
		Resolución de conflictos
		Valores

Nota: Zabalza, 2022.

La era de la IA generativa exige un replanteo profundo de la evaluación, al repensar cómo y por qué se evalúa a los estudiantes. Las tareas de evaluación deben estar diseñadas para alinearse con el proceso de aprendizaje y dejan de ser instantáneas evidenciadas en la producción puntual de un desempeño como resultado final (Lodge et al, 2023).

Para Cope, Kalantzis y Sears Smith (2021, p. 5)

La evaluación es quizás el área de oportunidad más importante que ofrece la inteligencia artificial para un cambio transformador en la educación. Sin embargo, esto no es evaluación en sus formas convencionalmente entendidas. La evaluación basada en IA utiliza artefactos y procesos dramáticamente diferentes de las evaluaciones tradicionales... De hecho, la IA podría significar el abandono y el reemplazo de las evaluaciones tradicionales, y con esto una transformación en los procesos de educación.

Swiecki et al (2022) propone una lista de usos de la IA generativa con un profundo poder transformador de la evaluación, tal como la conocemos hasta el momento. Diseño de evaluaciones automatizadas para obtener evidencia que respalde las afirmaciones sobre el aprendizaje, retroalimentación y evaluación entre pares asistida por IA y evaluación automatizada de la escritura de los estudiantes. Técnicas IA que brindan una visión más continua del desempeño y, por lo tanto, información sobre el proceso de aprendizaje, junto con técnicas que ajustan la tarea a las habilidades del estudiante, brindándole experiencias de evaluación personalizadas. La posibilidad de crear escenarios auténticos de aprendizaje mediado por simuladores, constituyen algunas de las opciones que hoy ofrecen las IA para cambiar los modos y fines de la evaluación.

Sin embargo, y más allá de las posibilidades técnicas, sigue siendo el factor pedagógico el punto de inflexión a la hora de resignificar la evaluación en tiempos de IA generativa. Consideraciones como diferenciar entre evaluar y medir el aprendizaje, considerar la compleja y variadas formas de aprendizajes, descritos en la tabla anterior; la autoevaluación y evaluación de pares para fomentar la autorreflexión entre los estudiantes sobre su propio trabajo, hasta la incapacidad de las IA generativa para reconocer, y mucho menos evaluar, resultados creativos al abordar una tarea, que destacan por su originalidad y que serían percibidos y valorados por un observador humano, constituyen ejemplo de lo que Swiecki et al (2022, p. 7) sentencia al afirmar que “existen aspectos del aprendizaje que siguen siendo perceptibles para los humanos pero no para las máquinas”.

Por tanto, resulta ineludible pensar qué, cómo y para qué evaluamos, de forma que la incorporación de estas tecnologías se alinea con los objetivos pedagógicos, abordando con determinación temas como la integridad académica, el uso ético y el desarrollo de habilidades de pensamiento crítico.

Autores como Perkins et al (2024) sostienen que los docentes necesitan herramientas para ayudar a los estudiantes a interactuar con la IA generativa de manera apropiada y ética; a la vez, las instituciones requieren mantener un enfoque estandarizado para su uso en situación de evaluación. Proponen la Escala de

Evaluación de IA (AIAS por sus siglas en inglés), consecuencia de un cambio en la percepción de la IA en el contexto de la educación universitarias australianas, que va en línea con la taxonomía descrita por Lodge (2021).

Añade que se comenzó categorizando el uso de IA en "Sin IA", "Alguna IA" y "IA completa", hasta perfeccionar una escala que recoge los aspectos más matizados del uso de la IA generativa, que permite cierto grado de discreción por parte de profesores y estudiantes y proporciona un sistema andamiado para integrarlas a la evaluación.

Tabla 3.

Escala de evaluación de la IA (AIAS) y actividades de evaluación sugeridas por nivel.

1	Sin IA	La evaluación se realiza completamente sin ayuda de IA, en ningún momento de la elaboración. Este nivel garantiza que los estudiantes confíen únicamente en sus conocimientos, comprensión y habilidades.	Discusiones, debates u otras formas orales de evaluación sin tecnología Ideación sin tecnología, trabajo individual o grupal en clase. Exámenes, sesiones de preguntas y respuestas, o formativos. Debates
2	Generar y estructurar ideas asistencia de IA	La IA se puede utilizar en la evaluación para generar ideas, crear estructuras y generar ideas para mejorar el trabajo con. No se permite contenido de IA en el envío final	Lluvia de ideas colaborativa. Esquemas estructurales: los estudiantes pueden usar IA para crear esquemas estructurados de su trabajo Asistencia en investigación, sugerir temas y fuentes de información
3	Edición asistida por IA	La IA se puede utilizar para mejorar la claridad o la calidad del trabajo creado por los estudiantes para mejorar el resultado final, pero no se puede crear contenido nuevo utilizando la IA. Se puede utilizar IA, pero su trabajo original sin contenido de IA debe incluirse en un apéndice	Gramática, puntuación y ortografía. Elección de palabras, sinónimos y reemplazo para mejorar la redacción. Ediciones estructurales para construir oraciones claras y coherentes. Edición de imágenes originales.
4	Tarea con IA y terminación humana de la evaluación	La IA se utiliza para completar ciertos elementos de la tarea, y los estudiantes brindan debates o comentarios sobre el contenido generado por la IA. Este nivel requiere un compromiso crítico con el contenido generado por IA y la evaluación de su resultado. Utilizará IA para completar tareas específicas en su evaluación. Se debe citar cualquier contenido creado por IA.	Generación directa de IA de datos, publicaciones y narrativas como base para generar sus propios productos originales. Análisis comparativo de semejanzas y diferencias de producciones humanas y generada por IA. Evaluación crítica para identificar sesgos e imprecisions,
5F	IA completa	La IA debe utilizarse como "copiloto" para cumplir los requisitos de la evaluación, permitiendo un enfoque colaborativo con la IA y mejorando la creatividad. Puede utilizar IA a lo largo de su evaluación para respaldar su propio trabajo y no es necesario especificar qué contenido se genera mediante IA.	Cocreación con iteran activamente sobre el contenido. Exploración de una amplia gama de temas, ideas, soluciones y estilos utilizando diversas herramientas IA. Retroalimentación de su trabajo en tiempo real hasta llegar al resultado final.

Nota: Perkins et al (2024)

Las investigaciones coinciden en destacar la importancia de ofrecer directrices concretas y claras a los estudiantes sobre lo que se espera de ellos, en especial en relación con el uso de la IA generativa en tareas de evaluación (Furze, 2024). La escala propone un sistema flexible para integrar de manera progresiva el uso de la IA generativa. Ayuda a los docentes a ajustar sus evaluaciones, decidiendo su uso en alineación con los objetivos de aprendizaje. A los estudiantes, la escala ofrece un marco abierto de cómo y cuándo usar la IA generativa, favoreciendo el desarrollo de una actitud de integridad académica.

La escala proporciona un enfoque estructurado que facilita la progresividad, donde cada nivel especifica el alcance del uso permitido de IA generativa y la responsabilidad de los estudiantes. La progresión permite adaptar la inclusión de la IA según las habilidades logradas por los estudiantes tanto en el uso como en la evaluación crítica de los resultados generados por las herramientas, que respalden la comprensión de los estudiantes sobre cómo las herramientas pueden usarse de manera efectiva y ética para respaldar las evaluaciones.

Herramientas como la descrita seguirán surgiendo en la medida en que los sistemas incluyen la IA generativa en la evaluación de los aprendizajes, con importantes demandas de actualización y formación continua por parte de los docentes,

Conclusiones

Los debates sobre la evaluación con el uso de IA exigen transparencia en las formas de inclusión, a la vez que se debe tener un claro reconocimiento de las limitaciones de la tecnología, ya que hay aspectos que quizás nunca pueda ser evaluados por la IA, siendo el papel del docente irremplazable. Proporcionar retroalimentación y mejorar el aprendizaje seguirán siendo órbita de la intencionalidad pedagógica, más allá de que sea asistido por IA generativa.

Los cambios desafían no sólo cómo se realizan las evaluaciones, sino que también implican un cambio en la concepción de su propósito y significado. Aunque la IA puede introducir innovaciones en las prácticas educativas, también plantea interrogantes importantes que educadores, formuladores de políticas y estudiantes deben enfrentar en términos de la alfabetización, pensamiento crítico y principios éticos.

A la pregunta ¿Puede la IA reemplazar a los docentes en la tarea de evaluar los aprendizajes de los estudiantes en todas sus dimensiones?, CHAT GPT4 responde:

La IA puede mejorar y apoyar ciertos aspectos de la evaluación, especialmente en tareas cuantificables y procesamiento de grandes cantidades de información. Sin embargo, no puede reemplazar a los docentes en la evaluación integral de los aprendizajes, ya que existen dimensiones como el pensamiento crítico, la creatividad, la ética y la capacidad de resolución de problemas que requieren una interpretación humana matizada. Los docentes aportan empatía, intuición y comprensión del contexto, aspectos clave para guiar el desarrollo personal de los estudiantes, lo que hace que la combinación de la inteligencia artificial y humana sea esencial para una evaluación completa y efectiva. (OpenAI, 2024).

Eso demuestra que los docentes tenemos un espacio privilegiado para reflexionar y generar transformaciones sustanciales en la práctica docente a la luz de los desafíos y oportunidades que ofrece la IA generativa.

Referencias

Bigg, J. (2006). *Calidad del aprendizaje universitario. Cómo aprenden los estudiantes*. Nacea.

- Caballero, E. (2022). Audiencias educativas. Impacto del entorno mediático en la realidad virtual. *AULA Revista de Humanidades y Ciencias Sociales*, 68(2), 9-21. <https://doi.org/10.33413/aulahcs.2022.68i1.203>
- Cope, B., Kalantzis, M., & Searsmith, D. (2020). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. *Educational Philosophy and Theory*, 53(12), 1229–1245. <https://doi.org/10.1080/00131857.2020.1728732>
- Fullan, M., & Langworthy, M. (2014). *A rich seam: How new pedagogies find deep learning*. Pearson. https://www.learning1to1.net/resource_redirect/downloads/sites/89909/themes/1506762/downloads/TAbDXr4vSTSqVDAFKyRs_3897.Rich_Seam_web.pdf
- Furze, L. (2024). Crossing the digital divide: Strategies for using GenAI in diverse classrooms. *Australian TAFE Teacher*, 58(2), 17-19. <https://search.informit.org/doi/abs/10.3316/informit.T2024072500012490591151018>
- Gobierno de España. (2024). *Guía sobre el uso de Inteligencia Artificial en el ámbito de la educación*. Instituto Nacional y de Formación de Profesorado. https://code.intef.es/wp-content/uploads/2024/07/Gu%C3%ADa-sobre-el-uso-de-la-IA-en-el-%C3%A1mbito-educativo-INTEF_2024.pdfv
- Lodge, J, Howard, S & Broadbent, J. (1 de mayo de 2023) *Assessment redesign for generative AI: A taxonomy of options and their viability*. LinkedIn <https://www.linkedin.com/pulse/assessment-redesign-generative-ai-taxonomy-options-viability-lodge/>
- OpenAI. (2024). *ChatGPT* (Versión GPT-4). <https://chat.openai.com/>
- Pozo, I. (2001). *Aprendices y maestros. La nueva cultura del aprendizaje*. Alianza.
- Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice*, 21(06). <http://open-publishing.org/journals/index.php/jutlp/article/download/810/769>
- Ritchhart, R., Church, M., & Morrison, K. (2014). *Hacer visible el pensamiento*. Paidós.
- Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, <https://doi.org/10.1016/j.caeai.2022.100075>
- Zabalza, M. (2022) *Coreografías Didácticas en Educación Superior*. Narcea.