

Evaluación sumativa en el colegio en la era de la Inteligencia Artificial (IA)

2023 vs. 2.0

Sobre el autor

Ronald Knust Graichen (holandés): Máster en Ciencias de Educación; especialista en tecnología educativa, internacionalización curricular, y en educación basada en competencias y resultados de aprendizaje, Universidad AERES, Holanda.

www.ronaldknustgraichen.com; ronaldknust@gmail.com

Grupo meta

Docentes de colegios.

Resumen

La evaluación sumativa¹ en la era de la inteligencia artificial (IA) es un desafío para los docentes en el colegio. Los modelos de IA como Bard y ChatGPT pueden generar texto que es indistinguible del escrito por humanos. Esto plantea el problema del plagio, ya que los alumnos pueden usar estos modelos para generar respuestas a exámenes y tareas sin realmente aprender el material.

El artículo ofrece una serie de estrategias que los profesores pueden usar para mitigar el riesgo de plagio en las evaluaciones sumativas.

Al utilizar una combinación de estas estrategias, los profesores pueden garantizar que sus evaluaciones midan con alta precisión el conocimiento y las habilidades de los alumnos.

Mitigación de riesgo de mal uso de ChatGPT o Bard en evaluaciones sumativas

En tiempos de ChatGPT y Bard, evaluar a los alumnos puede presentar desafíos adicionales debido a la posibilidad de copiar y hacer trampa utilizando estas herramientas de generación de texto.

Por un lado, la sociedad quiere la garantía de que los graduados del colegio tengan las habilidades adecuadas sin recurrir a ChatGPT, Bard u otros modelos de Inteligencia Artificial, pero por otro lado, el docente también debe enseñar al alumno cómo usar la IA en la vida cotidiana o laboral. El uso de software de detección tampoco tiene mucho sentido, ya que es muy laborioso para el docente y el alumno siempre encontrará nuevas formas de usar herramientas de IA que no pueden ser detectadas por el software de detección.

Antes de describir diversas estrategias para mitigar el uso inapropiado de IA en evaluaciones sumativas, se presentan algunas consideraciones sobre el uso de ChatGPT o Bard en relación con el plagio y la trampa en las evaluaciones:

- Definición de plagio: Según ChatGPT, el plagio se define como el acto de utilizar el trabajo o las ideas de otra persona sin dar el debido crédito a su autoría.
- Utilizar el contenido generado por ChatGPT o Bard sin citar adecuadamente las fuentes sería considerado plagio.

¹ La evaluación sumativa se realiza al final de un curso o unidad de estudio para determinar el nivel de aprendizaje de los alumnos. La evaluación formativa, por otro lado, se realiza durante el curso o unidad de estudio para proporcionar a los alumnos retroalimentación sobre su progreso y ayudarlos a aprender.

- **Responsabilidad académica:** Los alumnos deben asumir la responsabilidad de su propio aprendizaje y realizar sus tareas y evaluaciones de manera honesta. El uso de ChatGPT o Bard para generar respuestas completas sin esfuerzo propio sería considerado hacer trampa.
- **Deshonestidad académica:** La deshonestidad académica, como el plagio o la trampa en las evaluaciones, es una preocupación importante para los educadores.
- El uso de herramientas de inteligencia artificial para completar tareas o evaluaciones sin el esfuerzo y la comprensión adecuados no es ético y puede tener consecuencias negativas para el alumno.
- Utilizar estas herramientas para plagiar contenido puede llevar a ser detectado y enfrentar consecuencias académicas.² Por ejemplo, el alumno puede usar IA para
 - generar respuestas de exámenes o ensayos que son idénticos o similares a la información disponible en línea.
 - escribir diferentes tipos de contenido creativo, como poemas, historias o guiones, que pueden utilizar en sus tareas.

En resumen, los alumnos deben utilizar ChatGPT o Bard de manera responsable y ética. Estas herramientas pueden ser útiles para el aprendizaje, pero no deben ser utilizadas para hacer plagio o hacer trampa en las evaluaciones sumativas de una asignatura. Es importante que los alumnos asuman la responsabilidad de su propio aprendizaje y realicen sus tareas y evaluaciones de manera honesta y ética.

Estrategias que el docente puede utilizar para evaluar a los alumnos de manera efectiva.

En tiempos de ChatGPT y Bard, evaluar si el alumno ha logrado un resultado de aprendizaje esperado puede requerir enfoques adicionales.

A continuación, se presentan algunas estrategias que el docente puede utilizar para evaluar a los alumnos de manera efectiva:

1. *Evaluación auténtica:* Diseñar evaluaciones que reflejen situaciones reales y significativas que los alumnos puedan enfrentar en su vida profesional o personal. Esto puede incluir proyectos prácticos, estudios de casos o simulaciones que requieran la aplicación del conocimiento en contextos reales. Diseñar evaluaciones que requieran que los alumnos apliquen el conocimiento en situaciones prácticas o reales. Esto puede incluir la resolución de problemas, la realización de proyectos o la toma de decisiones basadas en escenarios específicos. Al evaluar la aplicación del conocimiento, se dificulta el uso de ChatGPT o Bard para simplemente copiar y pegar respuestas.

2. *Evaluación de competencias de orden superior:* En lugar de evaluar solo el conocimiento factual, enfocarse en evaluar competencias de orden superior como el pensamiento crítico, el análisis, la síntesis, la comunicación efectiva y la creatividad. Estas competencias son más difíciles de replicar utilizando herramientas de generación de texto y requieren una comprensión profunda y personal del contenido.

3. *Evaluación supervisada:* Realizar evaluaciones en un entorno supervisado, como exámenes en clase o presentaciones orales. Esto permite a los profesores observar directamente el desempeño de los alumnos y detectar cualquier intento de copia o trampa.

² Métodos de detección y prevención de plagio: <https://blog.classpoint.io/es/9-mejores-detectores-de-ia-gratuitos-para-profesores-recomendados-por-profesores-y-profesionales/>

4. *Evaluación personalizada*: Diseñar evaluaciones que sean únicas para cada alumno. Esto puede incluir asignaciones o proyectos individuales en los que los alumnos tengan que demostrar su comprensión y habilidades de manera personalizada, lo que dificulta la copia y la trampa.

5. *Variedad de métodos de evaluación*: Al igual que en la evaluación formativa, los profesores pueden utilizar una combinación de métodos de evaluación para obtener una imagen más completa del desempeño de los alumnos. Algunos ejemplos incluyen proyectos en clase, presentaciones orales, capacidad de trabajo en equipo, participación en clase, pensamiento crítico en debates en vivo, y resolución de problemas.

6. *Evaluación basada en evidencia*: Solicitar a los alumnos que proporcionen evidencia o argumentos respaldados por fuentes confiables para respaldar sus respuestas. Esto fomenta la comprensión crítica del tema, lo que dificulta el uso de herramientas de generación de texto para obtener respuestas sin fundamento.

7. *Uso permitido de IA*. En cualquier evaluación o tarea en la que el uso de herramientas de IA generativa esté explícitamente permitido, los alumnos deben confirmar cómo se han utilizado las herramientas de IA generativa (o no). Algunos ejemplos podrían ser (University of Birmingham, 2023):

- Requerir que los alumnos incluyan una declaración predefinida que indique explícitamente si han usado o no herramientas de IA generativa.
- Pedirles a los alumnos que compartan los “prompts” utilizados, los resultados o las modificaciones.
- Requerir que los alumnos carguen un componente reflexivo que detalle cómo se ha utilizado la IA generativa y su experiencia al interactuar con ella.

8. *Uso de herramientas antiplagio*: Utilizar herramientas antiplagio para detectar posibles copias o fuentes no citadas en los trabajos de los alumnos. Estas herramientas pueden ayudar a identificar cualquier intento de copia y desalentar a los alumnos de hacer trampa. Sin embargo, las herramientas antiplagio no siempre son precisas. Pueden detectar falsos positivos, es decir, trabajos que no son plagiados pero que son identificados como tales por la herramienta. Esto puede ser frustrante para los profesores y los alumnos.

En cuanto al proceso de aprendizaje, las herramientas antiplagio pueden tener un impacto negativo. Pueden hacer que los alumnos se sientan como si estuvieran siendo espiados. Esto puede crear un ambiente de desconfianza en el aula.

En cuanto a la carga de trabajo del profesor, las herramientas antiplagio pueden aumentarla. Los profesores tienen que dedicar tiempo a aprender a usar las herramientas, a interpretar los resultados y a tomar medidas disciplinarias contra los alumnos que han plagiado. Esto puede ser una carga adicional para los profesores que ya están sobrecargados de trabajo.

Lista de chequeo de Evaluación Sumativa resistente al copiado y la trampa

A partir del texto anterior, se puede deducir que con la llegada de la IA será necesario que el docente tenga que rediseñar algunas de sus evaluaciones sumativas para resistir al uso indebido de instrumentos de IA como Bard o ChatGPT por parte de los alumnos. A continuación, se presentan algunos criterios con los que el docente puede verificar si la evaluación diseñada por él está

preparada para resistir al uso indebido de la IA y puede medir si el alumno puede demostrar el resultado de aprendizaje esperado:

- *Contextualización y Aplicación Real:* Asegurarse de que las preguntas y tareas de la evaluación estén ancladas en situaciones reales y aplicables. De esta manera, será más difícil para los alumnos confiar únicamente en la IA para generar respuestas sin entender el contexto.
- *Preguntas Abiertas y Reflexivas:* Diseñar preguntas que requieran análisis crítico, síntesis y pensamiento reflexivo por parte del alumno. Esto dificultará que la IA genere respuestas completas y precisas sin el aporte genuino del alumno.
- *Uso de Escenarios Complejos:* Presentar escenarios complejos o problemas multifacéticos que requieran una comprensión profunda y habilidades de resolución de problemas que no puedan ser replicadas fácilmente por la IA.
- *Presentación de Argumentos Propios:* Solicitar a los alumnos que presenten argumentos o análisis originales respaldados por evidencia, lo que hace que sea más difícil para la IA imitar un razonamiento humano único.
- *Evaluación en Etapas:* Dividir la evaluación en etapas o partes sucesivas, donde el progreso y la coherencia sean evidentes. Esto dificultará que los alumnos dependan exclusivamente de la IA para generar respuestas en todas las partes.
- *Entrega Presencial o Supervisada:* Si es posible, llevar a cabo la evaluación en entornos presenciales o bajo supervisión, lo que reducirá la posibilidad de que los alumnos utilicen la IA de manera inapropiada.
- *Variedad en el Tipo de Preguntas:* Utilizar una variedad de tipos de preguntas, incluyendo opciones múltiples, verdadero/falso, completar espacios en blanco, y preguntas de desarrollo, para hacer que la evaluación sea más compleja y diversa.

Ejemplos de evaluaciones sumativas resistentes al mal uso de IA

A continuación, se presentan algunas formas de evaluación que sean resistentes al mal uso de IA para el resultado de aprendizaje "*Comer comida sana*". Estas estrategias buscan evaluar la comprensión profunda, la aplicación práctica y las habilidades de comunicación, lo que dificulta el uso de herramientas automatizadas para el engaño.

- Cree un plan de comidas para una semana que incluya una variedad de alimentos saludables. Esta tarea requiere que los alumnos tengan un conocimiento profundo de los diferentes grupos de alimentos y como combinarlos para crear una dieta equilibrada. ChatGPT o Bard no podría generar un plan de comidas que cumpla con todos los requisitos nutricionales.
- Investigue los beneficios de comer alimentos saludables y escriba un ensayo sobre sus hallazgos. Esta tarea requiere que los alumnos sean capaces de identificar y analizar información de una variedad de fuentes. ChatGPT o Bard no podría escribir un ensayo que sea bien investigado y bien escrito.
- Cree un juego educativo sobre la nutrición. Esta tarea requiere que los alumnos sean capaces de comunicar información compleja de una manera clara y atractiva. ChatGPT o Bard no podría crear un juego que sea divertido e informativo.

Conclusiones

Para evaluar a los alumnos en tiempos de ChatGPT y Bard, se pueden utilizar estrategias como la evaluación auténtica, la evaluación basada en habilidades, la evaluación supervisada, la evaluación personalizada, y la evaluación mixta. Estas estrategias ayudan a mitigar los riesgos de copia y trampa y promueven una evaluación más efectiva y justa. Es importante recordar que ninguna estrategia es infalible, pero al combinar diferentes enfoques de evaluación, se puede reducir la posibilidad de que los alumnos utilicen ChatGPT o Bard para copiar y hacer trampa. Además, es fundamental comunicar claramente las expectativas y los criterios de evaluación a los alumnos para fomentar la integridad académica.

Referencias

- Anderman, E. (2015, May 20). *Students cheat for good grades. Why not make the classroom about learning and not testing?* The Conversation. <https://theconversation.com/students-cheat-for-good-grades-why-not-make-the-classroom-about-learning-and-not-testing-39556>
- Mollick, E., & Mollick, L. (2023). Assigning AI: Seven approaches for students, with prompts. ArXiv. <https://arxiv.org/abs/2306.10052>
- Lee, J., & Soylu, M. Y. (2023). ChatGPT and Assessment in Higher Education. Retrieved from https://c21u.gatech.edu/sites/default/files/publication/2023/03/C21U%20ChatGPT%20White%20Paper_Final.pdf
- OpenAI. (2023). ChatGPT. Recuperado de <https://openai.com/blog/chatgpt/>
- University of Birmingham. (2023). Generative Artificial Intelligence and its Role Within Teaching, Learning and Assessment. Recuperado de <https://www.birmingham.ac.uk/university/hefi/gai/index.aspx>
- Torrey Trust. (2023, August 2). Essential Considerations for Addressing the Possibility of AI-Driven Cheating, Part 1. Retrieved from <https://www.facultyfocus.com/articles/teaching-with-technology-articles/essential-considerations-for-addressing-the-possibility-of-ai-driven-cheating-part-1/>
Part 2. Retrieved from <https://www.facultyfocus.com/articles/teaching-with-technology-articles/essential-considerations-for-addressing-the-possibility-of-ai-driven-cheating-part-2/>