



¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

Felipe I. Porflitt Becerra^{1,3,4,5}, Ricardo Rosas^{2,3} y Omar Valencia³

¹ Facultad de Educación. Pontificia Universidad Católica de Chile.

² Escuela de Psicología. Facultad Ciencias Sociales. Pontificia Universidad Católica de Chile.

³ Centro de Desarrollo de Tecnologías de Inclusión (CEDETi-UC).
Escuela de Psicología. Pontificia Universidad Católica de Chile.

⁴ Centro de Estudios de Políticas y Prácticas en Educación CEPPE UC

⁵ Cape Horn International Center (ANID/BASAL FB210018)

Resumen

Este estudio explora el uso de inteligencia artificial (IA) para analizar la creatividad en estudiantes universitarios/as chilenos/as, mediante la técnica *Think Aloud* aplicada a una submuestra de 13 personas pertenecientes a un estudio más amplio. En el marco de un diseño mixto, se emplearon herramientas como TurboScribe para transcribir audios, y ChatGPT para realizar análisis de contenido basados en la definición clásica de creatividad de Torrance (1966), evaluando fluidez, flexibilidad, originalidad y elaboración. El análisis se validó con la opinión de académicos/as y profesionales que compararon dos perfiles extremos (el más y el menos creativo), y por un test de creatividad, obteniendo resultados concordantes con los generados por la IA. Los hallazgos indican que, aunque la IA no reemplaza la creación humana, puede desempeñar un papel útil en la sistematización y análisis de datos cualitativos, al reducir carga de trabajo y aportar en la comprensión de factores complejos como la creatividad. Además, se destaca la posibilidad de aplicar principios de trazabilidad ética, como el uso de cadenas de bloques en contextos de investigación en ciencias sociales y/o educativa. Finalmente, se discuten limitaciones vinculadas a las plataformas de IA y a la necesidad de validación humana en procesos interpretativos, especialmente en estudios como este.

Palabras clave: *Inteligencia artificial. Creatividad. Análisis de contenido. Think Aloud. Educación superior.*





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

Introducción

El potencial innovador que tiene el uso de inteligencia artificial (IA) en el mundo de las ciencias, se ha visto favorecido por los constantes cambios en sus códigos, y marca un camino desafiante respecto a lo que esto conlleva para el futuro (Odeh, 2024). Existe un grado de saturación relevante en los modelos de lenguaje de IA, los que funcionan más bien de manera aislada (se alimentan en sus mismos datos), donde, por ejemplo, se observan fallas en generar contenido científico original, alejándose de la capacidad humana de crear (Lozić & Štular, 2023). Es aquí donde los modelos de lenguaje a gran escala han tenido un rol preponderante, mostrando enfoques que actualmente amplían el entendimiento de los códigos, que anterior a esto, se realizaban de manera manual (Dunivin, 2025).

Asimismo, la IA muestra un número creciente de herramientas para las ciencias cognitivas, donde se aprecia que este tipo de tecnologías muestran mejoras en los procesos metodológicos de las investigaciones (Han et al., 2024). Un aspecto relevante respecto a esto es que actualmente existen varias herramientas de transcripción de audio a texto, desde las altamente complejas (p.ej., *Speechmatics*, *PyannoteAI*, *Echo Labs*), otras de código abierto, como *Google Recorder*, *Live Transcribe*, *Whisper*, *Braina*, *Bear File Converter*, o el transcriptor de *WhatsApp*, así como también, algunas donde se observan diferencias en el nivel de complejidad que pueden generar este tipo de transcripciones, en relación a sus versiones pagadas y no pagadas (p.ej., *TurboScribe*, *Sonix*, *Trint*, *Otter.ai*, *Jamie AI*, *Happy Scribe*, etc.). Estas características de la tecnología actual son relevantes al momento de analizar las interacciones humanas con las máquinas en investigación en ciencias, y tienen como antecedente que, pese a que algunas las personas pueden tener diferencias individuales respecto a cómo se desempeña cada día/a usuario/a-máquina, en la mayoría de los casos el uso de IA implica una carga menor de trabajo cuando se vincula al mundo educativo (Xue & Wang, 2022).

En este contexto, donde la IA muestra buenos resultados y herramientas en ciencias y educación, es que es razonable ampliar los límites que actualmente ofrece la literatura. Más allá de facilitar tareas operativas o analíticas, consideramos que la IA podría ofrecer análisis más profundos para las ciencias cognitivas, y en particular, para los aspectos de la





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

creatividad humana, que son debate tanto para estudios cuantitativos (*Israel-Fishelson & HersHKovitz, 2024*) como en estudios cualitativos (*Thabane et al., 2024*). El presente papel de investigación no busca indagar en aquel debate, sino, tomar la definición antigua de creatividad de Torrance (1966), y saber cómo la IA puede contribuir en estos aspectos, considerando que el pensamiento creativo es esencial para la resolución de problemas y la toma de decisiones (*Mukherjee et al., 2021*).

Un aspecto que se considera relevante en áreas ajenas a las ciencias sociales y la educación es el uso de cadenas de bloques (*blockchain*) que garantiza los aspectos éticos de la investigación en análisis de contenido en el mundo empresarial (*Chen et al., 2023*). Un blockchain se caracteriza por ser un modelo donde la toma de datos se realiza de manera lineal, y donde las personas a cargo de levantar la data van revisando cada uno de los pasos correspondientes a la toma de muestra, tal como es el caso de este estudio, donde a partir del uso de IA, se quiso ver si un modelo como este era posible de validar en ciencias sociales. Además, concuerda con Lixandru (2024), quien respalda a la IA como una herramienta útil y confiable para análisis de contenidos cualitativos levantados a través de grupos focales. Como complemento a estas dos investigaciones, el presente estudio buscó indagar en el análisis de contenido para 13 transcripciones realizadas con IA, levantadas con una técnica de recolección de datos cualitativa (*Think Aloud*) que formó parte de un diseño metodológico más amplio (cuantitativo con insumos cualitativos), y que respetó el uso de cadenas de bloques, así como también, el foco específico de investigar en creatividad, en este caso, en estudiantes que estuvieran cursando tercer año de distintas carreras de una universidad chilena.

Así, como se describirá posteriormente, utilizamos la definición clásica de Torrance (1966) de creatividad (vale decir, donde fluidez, flexibilidad, originalidad y elaboración son las categorías relevantes de análisis), para hacer interactuar dos IA, y posteriormente realizar una consulta a personas que se desenvuelven trabajando en la misma universidad para comparar los resultados.





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

El presente estudio pretende contestar las siguientes preguntas:

- a) ¿Puede la IA distinguir los aspectos de fluidez, flexibilidad y originalidad planteados por la teoría de creatividad de Torrance?
- b) ¿Hay consistencia entre el criterio de más y menos creativo de la IA con un grupo de jueces humanos?
- c) ¿Hay correlación entre el juicio de la IA respecto del grado de creatividad de las respuestas de los/as estudiantes y los resultados de un test estandarizado de creatividad?.

Método

Muestra

Se utilizó una submuestra de un estudio más amplio que consideró 253 estudiantes de pregrado (tercer año) de las carreras de Psicología, Ingeniería, Diseño y Educación. El estudio más amplio tiene por objetivo estudiar la relación entre Inteligencia fluida, cristalizada y creatividad. A la muestra completa de estudiantes se le aplicaron instrumentos para medir los tres constructos. Para efectos del presente estudio, resultaba de interés conocer los procesos de razonamiento subyacentes a la resolución de problemas complejos. Para ello, se seleccionó de manera aleatoria una muestra de 13 estudiantes de las distintas carreras, para someterlos a un proceso de “Think Aloud” de la prueba de inteligencia fluida y cristalizada.

Instrumentos

Para la evaluación de la inteligencia fluida y cristalizada se utilizó el test M10 (Rosas y Pizarro, 2023). Se trata de una prueba de 34 ítems que tiene un límite de tiempo de 10 minutos, que presenta un buen nivel de confiabilidad (*Alfa de Cronbach*= .738). Para evaluar la creatividad, se aplicaron seis subpruebas de las escalas de creatividad del test TEDIB (Rosas, 1996).





Procedimiento

Fase 1: Evaluación de estudiantes

En una primera instancia (octubre a diciembre de 2024) se aplicó el test M10 y las subpruebas de TEDIB a la muestra completa, y en enero de 2025, se volvió a aplicar la prueba M10 a una sub-muestra de 13 estudiantes. Esta segunda vez, se aplicó el test sin considerar límites de tiempo, pero con la condición de relatar en voz alta sus estrategias (*Think Aloud*). Un psicólogo experto realizó y acompañó esta toma de datos para el pensamiento en voz alta.

Fase 2: Análisis de datos con IA

Para el análisis de datos, las respuestas fueron grabadas en audio (.mp3), y cada uno de los 13 audios se insertó en la plataforma TurboScribe de manera separada. Esta plataforma, permite hacer hasta tres transcripciones diarias en su versión gratuita, cada una de un máximo de 30 minutos de duración, lo que calzó tanto con la toma de datos (no hubo más de tres casos diarios), así como con la duración de las respuestas.

En particular, TurboScribe tiene tres calidades para realizar transcripciones: a) guepardo, que es la más rápida pero que puede contener errores mayores, b) delfín, que es de calidad y velocidad media, y c), ballena, que es la más lenta, pero donde la probabilidad de cometer errores es la más baja. La plataforma funciona de manera autónoma, y permite extraer las transcripciones en .pdf, .docx., .txt., .csv, .srt., y .vtt. Además, en caso de utilizar alguna de las tres primeras (.pdf, .docx., y .txt) permite la inserción de marcas de tiempo. Asimismo, permite reconocer hasta ocho hablantes, o determinarlos de manera automática, y, además, obliga a seleccionar el idioma de la transcripción. Para el caso de las 13 transcripciones se realizó el mismo procedimiento, recibiendo de esta manera la misma información de entrada: a) la calidad más alta de transcripción (ballena), b) reconocer entre dos hablantes (participante y entrevistador), c) idioma español, y d), utilizar marcas de tiempo.

Existen distintas alternativas para descargar el o los archivos que se generan con las transcripciones. En el caso de este estudio, se optó por la alternativa más directa, que consiste en vincular la cuenta de TurboScribe con ChatGpt. De esta forma, no es necesario extraer las transcripciones y volver a cargarlas en la segunda IA. Para que esto suceda es necesario dar una orden directa a la IA, y ambas aplicaciones deben estar asociadas a la misma cuenta de Google.





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

Para analizar el contenido de las transcripciones se realizaron los siguientes pasos: a) se accedió a una cuenta de Google en el navegador Chrome. b) Se accedió a las cuentas de TurboScribe y ChatGpt en pestañas del mismo navegador. c) Se le pidió a TurboScribe que abriera una transcripción desde ChatGpt, lo que resultó en un nuevo chat. d) Se escribió el siguiente prompt en el chat que se abrió en ChatGpt: “quiero ver las transcripciones y después te pediré algunas cosas”. Una vez que mostró las transcripciones listadas con hipervínculos, se le dio un prompt más específico: e) “quiero que sepas que cada una de esas transcripciones es de un test que se llama M10 (34 ítems). Siempre hay un orador/a que va relatando estrategias para resolver los problemas de esos ítems, y hay que basarse solo en ese orador/a”. Posteriormente, se le dio una instrucción para analizar uno de los casos: “Podríamos probar con ThinkAloudCaso1 y hacer lo siguiente: a) contar el número de estrategias para cada ítem, b) hacer un análisis de contenido por ítem, c) resumir y analizar si la persona tendría un nivel básico, medio o avanzado en resolución de problemas, y d) contrastar con datos generales si son estrategias “creativas” considerando que son personas de 21 años de edad (*Torrance, 1966*).” Esta última condición fue un requerimiento del estudio. ChatGpt se aseguró de que el resultado fuese el requerido, por lo que corroboró con una pregunta previa a realizar los análisis: “¡Genial! Ahora que tengo el contenido de la transcripción de ThinkAloudCaso1, haré lo siguiente: (1) Contar el número de estrategias por cada ítem. Voy a extraer cada estrategia usada en los distintos ítems y enumerarlas. (2) Análisis de contenido por ítem. Voy a analizar cómo el/la participante resuelve cada problema, qué enfoques usa y si sigue patrones lógicos, reglas matemáticas, intuición visual o estrategias verbales. (3) Evaluar el nivel de resolución de problemas. Voy a clasificar su desempeño en básico, medio o avanzado, basándome en la complejidad y eficacia de sus estrategias. (4) Contrastar con creatividad en jóvenes de 21 años. Voy a comparar sus estrategias con modelos de resolución de problemas creativos, considerando aspectos como flexibilidad, originalidad y fluidez en el pensamiento. Voy a procesar estos puntos ahora y te daré los resultados en breve. La respuesta a ese chat fue: “gracias. Aquí estoy esperando”, y la inteligencia artificial no realizó el análisis hasta tener una retroalimentación.

Una vez ejecutado este paso, la IA entregó los resultados anidados en los cuatro puntos nombrados. Para el primer caso -conteo de estrategias por ítem-, el resultado se mostró con: una tabla con el número de ítem, el número de estrategias utilizadas, y una descripción breve





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

de la/s estrategia/s del/la participante, además de una frase de conclusión. Para el segundo punto -análisis de contenido por ítem-, se detalló que, en este caso, existieron estrategias de eliminación, visuales, matemáticas, verbales y por intuición (*ThinkAloudCaso1*). En el tercer caso -resolución de problemas-, categorizó al/la participante con un nivel medio-alto, según tres ámbitos: a) tiene un razonamiento estructurado y efectivo, b) no muestra muchas estrategias “out of the box” o alternativas creativas, y c), usa métodos eficientes y rápidos. Finalmente, para el cuarto caso -contrastar con creatividad en jóvenes de 21 años-, ChatGpt mostró cuatro medidas de “creatividad”, según a) fluidez (*cantidad de ideas*), b) flexibilidad (*variedad de enfoques*), c) originalidad (*novedad de soluciones*), y d) elaboración (*detalle en las respuestas*), más una conclusión general para este apartado en particular.

El proceso fue repetido con las otras 12 transcripciones, y mostró una regularidad en términos de respuestas, así como variabilidad en los perfiles cognitivos/creativos de las y los estudiantes.

Fase 3: Validación humana

Para poder llevar a cabo una validación del procesamiento de datos de la IA, se preguntó a 10 personas respecto a cómo categorizarían dos perfiles seleccionados como el “más creativo”, y el “menos creativo” según estos resultados parciales, considerando que la probabilidad de que 10 personas respondan con la misma dirección a estos dos perfiles según análisis de contenido es $\approx .00097$ (Cohen, 1960; Krippendorff, 2004). Como criterio de inclusión para participar de esta evaluación se determinó que las personas escogidas debían ser académicos/as, investigadores/as o profesionales en la universidad donde se realizó el estudio. Como se nombró, se tomaron los casos más extremos de los resultados, y se describieron dos perfiles según las respuestas que dieron las y los participantes, y los resultados analizados por ChatGpt. Esta parte del procedimiento se llevó a cabo en marzo de 2025. En específico, se solicitó a cada evaluador/a que leyera los perfiles de resolución de problemas de dos estudiantes. Se les explicó que este perfil se había realizado en base a las estrategias de expresión verbal utilizada para responder el test M10. Además, se les pidió que evaluaran el nivel de creatividad de los/as estudiantes utilizando una escala de Likert de 0 (*nada creativo*) a 4 (*muy creativo*)”.





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

Los resultados indican la presencia de coincidencia entre la evaluación realizada por humanos y la realizada por IA. Tal como se puede ver en la tabla 1 el/la primer/a estudiante categorizado por ChatGpt como el “más creativo” obtuvo una puntuación más alta ($M=3.2$; $DE=0.63$) que el segundo estudiante ($M=1.3$; $DE=0.48$), el cual fue categorizado por la IA como el “menos creativo” ($t(6)=8.14$, $p<0.001$), mostrando un tamaño de efecto igual a $d=2.57$ (Cohen, 1988). Esto indica la existencia de coincidencia entre lo realizado por ChatGpt y la evaluación realizada por humanos. Todos/as los/as evaluadores/as que participaron del proceso de validación categorizaron más creativo al estudiante correspondiente al primer perfil, lo que coincide con lo realizado por ChatGpt; por otra parte, incluso en aquellos casos donde la diferencia de puntaje fue menor (1 punto de diferencia entre los perfiles en el caso de los evaluadores 1 y 8), la dirección se mantuvo. Cabe mencionar que en ninguno de los dos casos se marcó la respuesta “0” de la escala, vale decir, ningún/a evaluador/a considera “nada creativo/a” a los/as estudiantes. Los resultados se muestran a continuación en la Tabla 1:

Tabla 1

Puntuación realizada por los evaluadores para cada uno de los perfiles.

N	Perfil 1	Perfil 2
Evaluador/a 1	3	2
Evaluador/a 2	3	1
Evaluador/a 3	3	1
Evaluador/a 4	3	1
Evaluador/a 5	3	1
Evaluador/a 6	2	1
Evaluador/a 7	4	2
Evaluador/a 8	3	2
Evaluador/a 9	4	1
Evaluador/a 10	4	1
Promedio	3,2	1,3
Desviación estándar	0,63	0,48





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

Fase 4: Validación de puntaje ChatGpt con puntaje en prueba de creatividad

Por último, los puntajes totales en creatividad calculado por ChatGpt se correlacionaron con los puntajes obtenidos por los/as 13 participantes seleccionados para la fase de “think aloud”. En la Tabla 2 se muestran las posiciones ordinales de cada participante según cada uno de los criterios:

Tabla 2

Posiciones relativas en ambas medidas de creatividad.

Ranking Según Prueba TEDIB	Ranking Según IA
1	1
2	2
3	3
4	4
6,5	5
6,5	6
6,5	7
6,5	8
9	9
10	10
11	11
12	12
13	13

Nota: Los puntajes TEDIB están rankeados según la suma de puntajes brutos en las tres pruebas de creatividad aplicadas. Los casos 5-8 comparten rango por tener el mismo puntaje. Se asigna rango promedio.

La correlación de Spearman entre ambas medidas fue igual a $r = .986$ ($p < 0.01$), la que puede ser considerada como muy alta y estadísticamente significativa.





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

Discusión

Estos resultados nos permiten concluir que la IA es capaz de analizar y puntuar de manera adecuada el contenido de las transcripciones que se presentaron para los 13 casos del estudio (*Think Aloud*), en términos de la creatividad. No obstante, a este primer acercamiento positivo, y considerando que podría haber mejoras en procedimientos como este, en el caso de esta investigación los pasos fueron rigurosos en términos de lo que se le pidió como resultado a la transcripción (*TurboScribe*), y también, en cómo se vinculó a ChatGpt para los análisis, lo que calza con lo propuesto para los bloques en cadena típicamente utilizados en el área empresarial (*Chen et al., 2023*), esta vez en ciencias sociales.

También consideramos que los procedimientos que aún se realizan de manera manual pueden ser generados en la actualidad por análisis de IA, concordando con Dunivin (2025) respecto a que la tecnología actual permite ampliar las fronteras de los campos de análisis, siempre y cuando la forma de llevar a cabo las interacciones humano-máquina sean las adecuadas, y así, efectivamente se vuelven desafiantes (*Odeh, 2024*).

Finalmente, según Han et al. (2024) este tipo de uso para investigación en ciencias sociales es sumamente útil. Nuestros hallazgos concuerdan con esto y, además, consideran que una validación humana es necesaria a lo menos en este período inicial de interacción de personas con IA, para aportar en la discusión respecto a si la máquina puede mermar el pensamiento humano. Por el contrario, consideramos razonable con estos datos plantear que el uso de IA en el procesamiento de datos cualitativos ahorra tiempo y recursos humanos.

Limitaciones

Futuras investigaciones podrían considerar que las IA tienen cambios de políticas de manera recurrente, y, además, consideran un aprendizaje veloz según la tecnología avanza, razón por la cual este estudio se enmarca específicamente en lo que fue posible llevar a cabo en marzo de 2025. Dentro del mismo punto, y en concordancia con los hallazgos de Lozić & Štular (2023), la creación humana aún no es posible de emular por la IA, sin embargo, se requiere de constante revisión respecto al uso de este tipo de tecnologías, dado que su velocidad de avance es muy alta.





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

Autores

Felipe I. Porflitt Becerra. Profesor Asistente. Facultad de Educación. Pontificia Universidad Católica de Chile. Investigador Asociado Cape Horn International Center (ANID/BASAL FB210018). Investigador Asociado Centro de Estudios de Políticas y Prácticas en Educación CEPPE UC. Miembro del Directorio del Centro de Desarrollo de Tecnologías de Inclusión CEDETi UC.

Ricardo Rosas. Profesor Titular. Escuela de Psicología. Facultad Ciencias Sociales. Pontificia Universidad Católica de Chile. Director del Centro de Desarrollo de Tecnologías de Inclusión CEDETi UC.

Omar Valencia. Asistente de Investigación. Centro de Desarrollo de Tecnologías de Inclusión CEDETi UC. Pontificia Universidad Católica de Chile.

Referencias

- Chen, J. (Elaine), Bao, F., Li, C., & Lin, Y. (2023). The Application and Ethics of Artificial Intelligence in Blockchain. *Journal of Global Information Management*, 31(7), 1–32. <https://doi.org/10.4018/JGIM.323656>
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. In *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Dunivin, Z. O. (2025). Scaling hermeneutics: a guide to qualitative coding with LLMs for reflexive content analysis. *EPJ Data Science*, 14(1), 28. <https://doi.org/10.1140/epjds/s13688-025-00548-8>





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

- Han, J., Childs, P. R. N., & Luo, J. (2024). Applications of artificial intelligence and cognitive science in design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 38, e6. <https://doi.org/10.1017/S0890060424000052>
- Israel-Fishelson, R., & HersHKovitz, A. (2024). Log-Based Analysis of Creativity in the Context of Computational Thinking. *Education Sciences*, 15(1), 3. <https://doi.org/10.3390/educsci15010003>
- Krippendorff, K. (2004). Reliability in Content Analysis. *Human Communication Research*, 30(3), 411–433. <https://doi.org/10.1111/j.1468-2958.2004.tb00738.x>
- LIXANDRU, I.-D. (2024). The Use of Artificial Intelligence for Qualitative Data Analysis: ChatGPT. *Informática Económica*, 28(1/2024), 57–67. <https://doi.org/10.24818/issn14531305/28.1.2024.05>
- Lozić, E., & Štular, B. (2023). Fluent but Not Factual: A Comparative Analysis of ChatGPT and Other AI Chatbots' Proficiency and Originality in Scientific Writing for Humanities. *Future Internet*, 15(10), 336. <https://doi.org/10.3390/fi15100336>
- Mukherjee, M., Law, P., & Baid, K. (2021). Emotional Intelligence, Social Intelligence, and Creativity in Early Adolescents in an Urban School Sample in Kolkata. *Indian Journal of Health & Wellbeing*, 12(3), 367–375. <https://openurl-ebsco-com.pucdechile.idm.oclc.org/contentitem/asn:153470035?sid=ebsco:plink:crawler&id=ebsco:asn:153470035&crl=c>
- Odeh, A. (2024). Exploring AI Innovations in Automated Software Source Code Generation: Progress, Hurdles, and Future Paths. *Informatica*, 48(8), 125–136. <https://doi.org/10.31449/inf.v48i8.5291>
- Rosas, R. (1996). Replicación del Modelo de Estructura de Inteligencia de Berlín en una Muestra de Estudiantes Chilenos. *Psykhē*, 5(1). <https://ojs.uc.cl/index.php/psykhe/article/view/20271>
- Rosas, R., & Pizarro, M. (2023). Reporte técnico de pruebas de habilidades cognitivas generales M10. CEDETest, Reportes Técnicos, Centro de Desarrollo de Tecnologías de Inclusión CEDETi UC. <https://www.cedeti.cl/wp-content/uploads/2023/03/Reporte-te%CC%81cnico-M10.pdf>





¿Puede la Inteligencia Artificial evaluar confiablemente la creatividad humana medida por los test?

- Thabane, A., McKechnie, T., Staibano, P., Arora, V., Calic, G., Busse, J. W., Parpia, S., & Bhandari, M. (2024). A consensus definition of creativity in surgery: A Delphi study protocol. *PLOS ONE*, 19(12), e0314445. <https://doi.org/10.1371/journal.pone.0314445>
- Torrance, E.P. (1966). *Torrance Tests of Creative Thinking*. Lexington, MA: Personnel Press
- Xue, Y., & Wang, Y. (2022). Artificial Intelligence for Education and Teaching. *Wireless Communications and Mobile Computing*, 2022(1), 1–10. <https://doi.org/10.1155/2022/4750018>

