



CONGRESO INTERNACIONAL VIRTUAL SOBRE LAS TECNOLOGÍAS DEL APRENDIZAJE Y DEL CONOCIMIENTO

Validación preliminar de
instrucciones para revisar
manuscritos académicos

**XVII
CIVTAC**

PONENTE:

Dr.(c) Emilio Chocobar Reyes

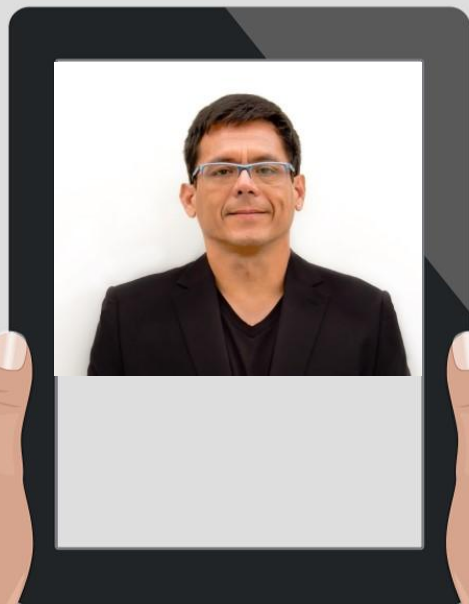
XVII CIVTAC



<https://orcid.org/0000-0002-5270-1162>



ECHOBAR@zegel.edu.pe



Síntesis curricular:

Emilio Chocobar Reyes

Doctorando en Educación, investigador principal y revisor externo de artículos en revistas científicas

Mis redes sociales son:



@emiliochocobar



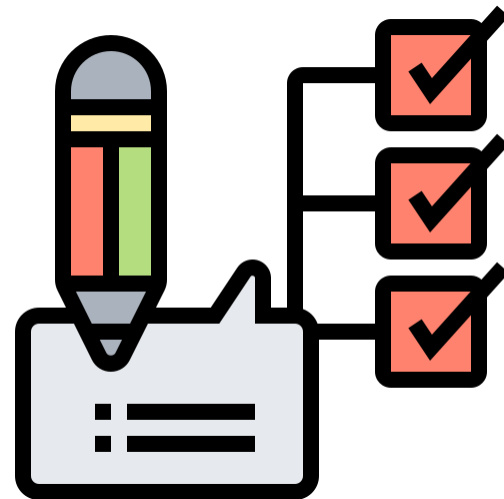
@emilio_chocobar



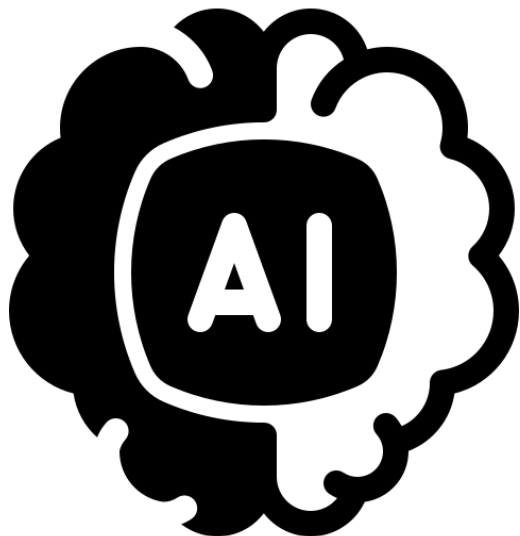
La revisión académica con IAG fortalece la calidad editorial al detectar debilidades estructurales, argumentativas y formales antes del envío a revistas científicas.

(Kaklauskienė & Kaklauskas, 2024; Negi & Srivastava, 2021)

Los procesos editoriales enfrentan rechazos, demoras y observaciones por fallas de estructura, metodología, novedad, claridad, discusión, autoría y adecuación editorial.



(Menon et al., 2022; Ganga et al., 2022; Migheli et al., 2021; Paul, 2024; Kumar et al., 2023; Patel, 2022).



IAG puede identificar inconsistencias y errores, pero requiere prompts organizados, supervisión crítica y controles frente a alucinaciones o referencias inexistentes.

(Liu et al., 2023; Cevallos et al., 2025; Damián et al., 2025; Chocobar et al., 2025; Gahrn, 2025; Guile, 2023; Cavalcanti de Arruda et al., 2022).

El estudio diseña y evalúa un prompt especializado para revisión crítica comparativa en ChatGPT y Perplexity, orientado a calidad editorial responsable.



(Chocobar & Barreda, 2025).

Diseñar y evaluar preliminarmente la funcionalidad de un prompt especializado para la revisión crítica de manuscritos académicos, mediante su aplicación comparativa en ChatGPT y Perplexity



Teoría de la Cognición Distribuida:
concibe la producción científica
como un proceso sociotécnico
donde interactúan investigadores,
herramientas digitales, prácticas y
entornos tecnológicos

(Gahrn, 2025; Guile, 2023; Cavalcanti de Arruda et al., 2022).

1

MÉTODO

Hipotético-deductivo

PARADIGMA

Positivista

ENFOQUE

Cuantitativo

TIPO

Exploratorio-aplicado

DISEÑO

No experimental,
validación preliminar

2

POBLACIÓN

Manuscritos
científicos
disponibles en
fuentes académicas

MUESTRA

46 artículos
científicos de acceso
abierto (2025).

Muestreo: No
probabilístico
intencional

3

TÉCNICA

Análisis documental,
búsqueda académica y
aplicación controlada de
instrucciones

HERRAMIENTA

Prompt 10 criterios:
Título, Resumen,
Introducción,
Metodología, Resultados,
Discusión-Conclusiones,
Bibliografía, Rigor,
Originalidad y Relevancia
Escala Likert: 0–3
Puntaje: 0-30

4

ANÁLISIS

Estadística
Descriptiva
Shapiro-Wilk
Wilcoxon
Correlación rango-
biserial
Bland–Altman

Reglas de decisión editorial: Rangos de puntaje para toma de decisiones, mínimo 0 y máximo 30.

27–30 Publicarse tal como está

22–26 Correcciones menores

15–21 Correcciones mayores

0–14 No publicable

Tabla 5: Estadísticos descriptivos de los puntajes otorgados por cada herramienta de IAG

Herramienta de IAG	n	Media	DE	Mediana	RIC	Mín.	Máx.
Perplexity Pro con Investigación Profunda	46	21.11	1.90	21	20–22	18	26
ChatGPT 5.5 Ampliado	46	24.85	4.33	26	21–29	17	30

Nota: DE = desviación estándar; RIC = rango intercuartílico. Elaboración propia (2026).

Tabla 6: *Distribución de decisiones editoriales generadas por el prompt especializado*

Decisión editorial	Perplexity Pro con Investigación Profunda	%	ChatGPT 5.5 Ampliado	%
Publicarse tal como está	0	0.0	20	43.5
Publicarse con correcciones menores	19	41.3	12	26.1
Publicarse con correcciones mayores	27	58.7	14	30.4
No publicable en la revista	0	0.0	0	0.0
Total	46	100.0	46	100.0

Nota: La tabla muestra la distribución porcentual de las decisiones tomadas por cada IAG. Elaboración propia (2026).

Tabla 8: *Comparación pareada de puntajes entre herramientas de IAG*

Comparación	Diferencia media	DE de la diferencia	Mediana de diferencia	IC 95%	W	P	Tamaño del efecto
ChatGPT – Perplexity	3.74	3.93	4.5	2.57 a 4.91	103.5	< .001	rrb = 0.80

Nota: rrb = correlación rango-biserial. Elaboración propia (2026).

Tabla 9: *Indicadores de asociación y acuerdo entre ChatGPT 5.5 Ampliado y Perplexity Pro con Investigación Profunda*

Indicador	Valor	p
Correlación de Spearman	.415	.004
Tau-b de Kendall	.329	.003
Kappa no ponderado	-.036	—
Kappa ponderado lineal	.035	—
Kappa ponderado cuadrático	.104	—
ICC(2,1), acuerdo absoluto	.192	—
ICC(2,k), promedio de herramientas	.322	—

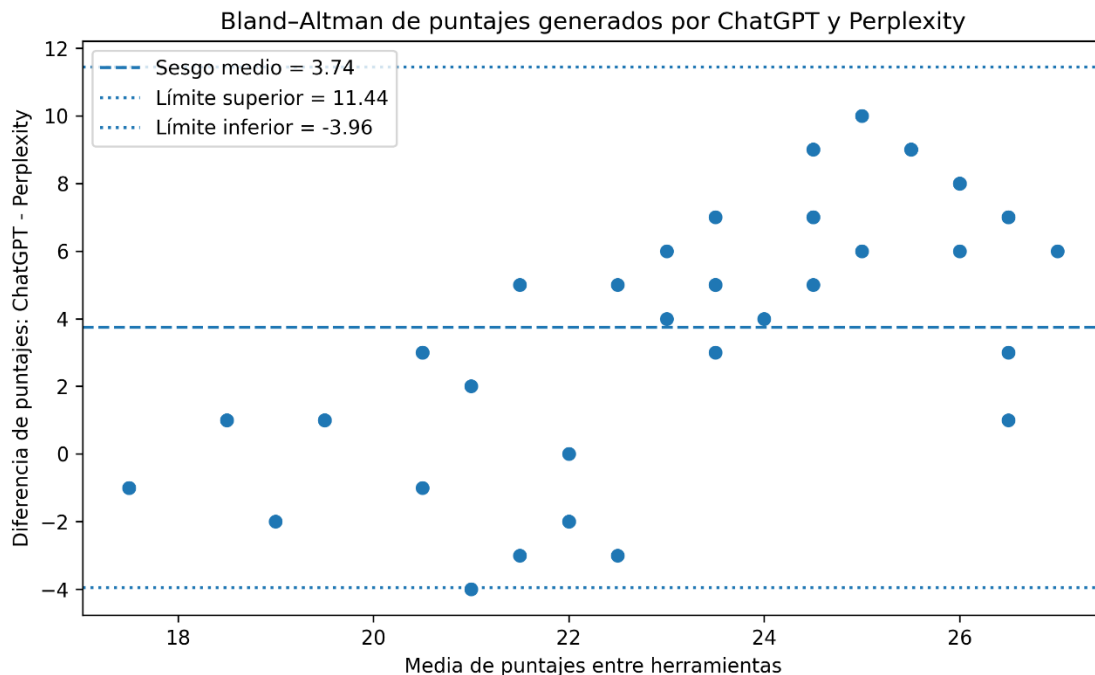
Nota: ICC = coeficiente de correlación intraclase. Los valores de kappa fueron calculados sobre las categorías editoriales ordinales. Elaboración propia (2026).

Tabla 10: *Análisis de Bland–Altman entre ChatGPT 5.5 Ampliado y Perplexity Pro con Investigación Profunda*

Indicador	Valor
Sesgo medio	3.74
DE de la diferencia	3.93
Límite inferior de acuerdo	-3.96
Límite superior de acuerdo	11.44
Diferencia mínima observada	-4.00
Diferencia máxima observada	10.00

Nota: La tabla muestra en análisis de Bland–Altman. Elaboración propia (2026).

Figura 1: *Gráfico de Bland–Altman de los puntajes generados por ChatGPT 5.5 Ampliado y Perplexity Pro con Investigación Profunda*



Nota: La figura presenta la media de los puntajes de ambas herramientas en el eje X y la diferencia ChatGPT – Perplexity en el eje Y. La línea central representa el sesgo medio y las líneas externas representan los límites de acuerdo.

- ➔ El prompt demostró funcionalidad preliminar.
- ➔ Los diez criterios permitieron evaluar integralmente.
- ➔ ChatGPT y Perplexity no evaluaron de manera equivalente.
- ➔ El prompt demostró potencial como herramienta preventiva antes del envío editorial.

- ➔ Aplicar el prompt en manuscritos y comparar versiones antes y después.
- ➔ Contrastar las evaluaciones de IAG con revisores humanos.
- ➔ Utilizar el prompt como herramienta preliminar y supervisada.

Gracias!

Alguna Pregunta?

Me puedes contactar a través de:

echocobar@zegel.edu.pe

echocobar@gmail.com

- Cabral, R., de Siqueira, I., de Araújo, A., Vieira, C., Santos, C., Gomes, E., Bezerra, P., da Costa, S. (2024). Preparation and validation of the instrument "QualiAPS digital—Brazil" for assessing digital health care in primary health care: a required tool. *Frontiers in Public Health*, 12, 1304148. <https://doi.org/10.3389/fpubh.2024.1304148>
- Cavalcanti de Arruda, G., Medeiros de Freitas, D., Soares, C., Nawrotek, K., & Miranda, B. (2022). The production of knowledge through religious and social media infrastructure: World making practices among Brazilian Pentecostals. *Popular Communication*, 20(3), 208–221. <https://doi.org/10.1080/15405702.2022.2074997>
- Cevallos, G., Ubillús, J., & Chocobar, E. (2025). Argumentos a favor de permitir o proibir el uso de la inteligencia artificial generativa por estudiantes. Una revisión sistemática. *European Public & Social Innovation Review*, 11, 1–30. <https://doi.org/10.31637/epsir-2026-2155>
- Chocobar, E., Ubillús, J. & Cevallos, G. (2025). ¿Sólo el ChatGPT? Diferentes herramientas de Inteligencia Artificial Generativa (GenAI) y sus usos: Una revisión de la literatura científica, 2021 – 2024. *Revista Científica De La UCSA*, 12(1), 70–85. <https://doi.org/10.18004/ucsa/2409-8752/2025.012.01.070>
- Chocobar, E. y Barreda, R. (2025). Estructuras metodológicas PICO y PRISMA 2020 en la elaboración de artículos de revisión sistemática: Lo que todo investigador debe conocer y dominar. *Ciencia Latina Revista Científica Multidisciplinar*, 9(1), 8525-8543. https://doi.org/10.37811/cl_rcm.v9i1.16491
- Costa, A., Rodrigues, F., Pitarma, R., & Ferreira, M. (2024). Design and validation of an instrument to evaluate the workgroup methodology in university students. *International Journal of Educational Research Open*, 7, 100351. <https://doi.org/10.1016/j.ijedro.2024.100351>
- Damián, M., Acosta, D., Delgado, I., & Rojas, G. (2025). Percepción de Estudiantes de Administración sobre la Utilidad de la Inteligencia Artificial en Artículos Científicos. *Revista Docentes 2.0*, 18(2), 107–126. <https://doi.org/10.37843/rtd.v18i2.678>
- da Silva, L., Silva, R., Pereira, M., Herdeiro, M., & Baldoni, A. (2024). Construction and validation of an instrument to identify barriers to implementing pharmaceutical care. *Exploratory Research in Clinical and Social Pharmacy*, 16, 100529. <https://doi.org/10.1016/j.rcsop.2024.100529>
- Eaton, S. (2023). Postplagiarism: Transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity*, 19, 23. <https://doi.org/10.1007/s40979-023-00144-1>
- Fei, W., Jingqi, T., Xiaonan, S., Xinying, S., Jun, L., Xingxing, M., & Yibo, W. (2025). Design and development of scales in primary care: Practical steps and statistical methods. *Chinese General Practice Journal*, 2(4), 100091. <https://doi.org/10.1016/j.cgpi.2025.100091>
- Ferrando, P., Morales, F., Casas, J., & Muñoz, J. (2025). Likert scales: A practical guide to design, construction and use. *Psicothema*, 37(4), 1–15. <https://doi.org/10.70478/psicothema.2025.37.24>
- Gahrn, R. (2025). Enlanguaged affordances in social practices: A critical rethinking of Gibson's approach to language. *Language & Communication*, 104, 121–130. <https://doi.org/10.1016/j.langcom.2025.07.004>