

# CARTA AL EDITOR

---

## Comentarios al artículo “Uso del ChatGPT en la mejora de la producción de textos”.

Estimado Sr./Sra. Editor/a,

he leído con interés el artículo “Uso del ChatGPT en la mejora de la producción de textos” publicado en el más reciente número de la Revista Peruana de Educación (Apolín, Lirión, Sánchez, & Aranda, 2024). Sin lugar a duda, un trabajo necesario dado su enfoque a imbricar estas herramientas en el futuro de la educación. A pesar de lo anterior, y con el afán de mejorar metodológicamente estudios similares, le propongo algunos comentarios, fundamentalmente en los apartados de métodos y resultados.

*Método:* En el artículo de Apolín Montes et al., (2024), se pueden observar algunos puntos de mejora en este apartado:

- “Diseño preexperimental”: según la información y los resultados presentados, este diseño no se corresponde con el diseño real de la investigación. Según la fuente citada por los autores (Chávez, Esparza, & Riosvelazco, 2020), este se ajusta más a un estudio cuasiexperimental, con dos grupos no equivalentes, uno control y otro experimental.
- Población y muestra: uno de los detalles más llamativos es que se presenta una muestra de 99 estudiantes repartidos en los dos grupos, pero en los resultados el grupo control tiene 39 y el experimental 35 participantes, o sea, 74 en total. Aunque se dan criterios de inclusión y exclusión, no queda claro si estos se aplicaron para seleccionar a los 99 o para descartar algunos otros y conformar una muestra final.
- Técnica: en el estudio se sostiene que la técnica utilizada fue la observación, pero la obtención de los datos se realiza a través de una rúbrica de evaluación enfocada a trabajos escritos. Se recuerda que esta técnica implica el registro sistemático objetivo, válido y confiable del comportamiento o conducta de los estudiantes durante las sesiones lo que no parece haberse desarrollado en el trabajo (Matos & Pasek, 2008). Según la información presentada, la técnica presentada podría ser el análisis documental (Martínez-Corona, Palacios-Almón, & Oliva-Garza, 2023) o también el análisis de contenido (Espín, 2002), aunque con la poca información presentada es difícil determinar.
- Instrumento: si bien la rúbrica es un instrumento poderoso para obtener datos objetivos en la evaluación de competencias, su uso requiere de un análisis de fiabilidad y validez (Cano, 2015). En este caso, en el estudio no queda claro el porqué de la asignación desigual de puntos o la ponderación de cada criterio. Se menciona que se validó por los docentes, pero no hay una información clara bajo qué criterios objetivos se obtuvo esta validez. Tampoco queda se explicitan las razones para no utilizar los criterios 1 y 2 que también se relacionan con la redacción académica, entre otros temas.

# CARTA AL EDITOR

---

*Resultados:* este apartado puede tener mejoras importantes, desde la presentación de los datos hasta el análisis estadístico, algunas de las cuales se detallan a continuación.

- Organizar la información de los datos “brutos” o procesados en tablas. Esto permitiría al lector tener una visualización más clara de los aportes del trabajo así como para valorar con precisión los resultados obtenidos (Lam, 2016; Branch & Villarreal, 2008). No se puede promediar frecuencias absolutas, considerando las limitaciones que esto implica para la interpretación de la media aritmética como estimador puntual de la media poblacional.
- En la tabla 1, muestro un ejemplo de cómo se podrían organizar los datos de frecuencias absolutas presentados por los autores del estudio analizado. Asimismo, se presenta la probabilidad de error tipo I en base a los supuestos del planteamiento de hipótesis de la prueba Ji-cuadrado de Pearson (McHugh, 2013). De estos resultados se interpreta que, en ninguno de los criterios, las frecuencias observadas fueron diferentes de las esperadas, o en otras palabras, las variables de grupo y resultados en la intervención, son independientes entre sí.

Tabla 1. Frecuencia de estudiantes según sus niveles de logro en la escala valorativa de cada criterio. (Adaptado de los resultados de Apolín Montes et al., 2024).

Criterio	Grupo	Nivel de logro				p
		Logrado	En proceso	Inicio	Total	
3	Control	12	20	7	39	0,2882
	Experimental	6	24	5	35	
4	Control	14	13	12	39	0,1587
	Experimental	7	19	9	35	
5	Control	0	38	1	39	NC *
	Experimental	0	35	0	35	

Nota: \* No se presenta la probabilidad de error tipo I porque un 50 % de las celdas tienen valores esperados menores a 1 (ver McHugh, 2013). Cálculos realizados con Epidat 3.1.

- Si en vez de lo anterior, desea realizar un análisis con el test T de Student para grupos independientes, promedie las puntuaciones finales en los trabajos entregados, específicamente en los criterios evaluados, tomando en cuenta sus ponderaciones. Con esto, plantee las hipótesis para grupos independientes y aplique el procedimiento tomando las consideraciones del caso (Hae-Young, 2019). En la tabla 2 quedaría un ejemplo de este procesamiento estadístico considerando los puntajes ponderados por criterios datos por los autores del artículo mencionado.

# CARTA AL EDITOR

---

Tabla 2. Puntuación total promedio de los criterios 3, 4 y 5, desviación estándar y resultados del test T de Student en la comparación de los dos grupos de estudio.

Grupo	N	Media	S	t	p
Control	39	3.74	1.27	0.107	0.915
Experimental	35	3.71	1.07		

- En la publicación no es recomendable presentar varias veces el mismo procedimiento de cálculo, debido a que esto le hace muy redundante. Se recomienda utilizar herramientas informáticas gratuitas que le ayuden en este proceso como Jamovi (<https://www.jamovi.org>) o Epidat (<https://www.sergas.es/Saude-publica/EPIDAT?idioma=es>), declarando las mismas en los procedimientos.

Sin otro tema que agregar, animarlos a que sigan publicando trabajos tan interesantes e importantes para el dominio de la educación en el contexto actual de las inteligencias artificiales.

Atentamente,

Dariel Díaz Arce

Orcid: <https://orcid.org/0000-0002-5952-9916>

Magíster en Innovación en Educación

Unidad Educativa Santana

Cuenca, Ecuador

## Referencias

- Apolín, L. A., Liri6n, I. R., S6nchez, R., & Aranda, L. A. (2024). Uso del ChatGPT en la mejora de la producci6n de textos acad6micos. *Revista Peruana de Educaci6n*, 6(12), 23-33. doi:<https://doi.org/10.37260/repe.v6n12.2>
- Branch, L., & Villarreal, D. (2008). Redacci6n de trabajos para publicaciones cient6ficas. *Ecolog6a Austral*, 18, 139-150. Obtenido de <http://www.scielo.org.ar/pdf/ecoaus/v18n1/v18n1a11.pdf>
- Cano, E. (2015). Las r6bricas como instrumento de evaluaci6n de competencias en educaci6n superior: ¿uso o abuso? *Profesorado. Revista de Curr6culum y Formaci6n del Profesorado*, 19(2), 265-280. Obtenido de <https://www.ugr.es/~recfpro/rev192COL2.pdf>

# CARTA AL EDITOR

---

- Chávez, S. M., Esparza, O. A., & Riosvelazco, L. (2020). Diseños preexperimentales y cuasiexperimentales aplicados a las ciencias sociales y a la educación. *Enseñanza e Investigación en Psicología*, 2(2), 167-178. Obtenido de <https://revistacneipne.org/index.php/cneip/article/view/76>
- Espín, J. V. (2002). El análisis de contenido: una técnica para explorar y sistematizar información. *XXI, Revista de Educación*, 4, 95-105. Obtenido de <https://rabida.uhu.es/dspace/bitstream/handle/10272/1913/b15141895.pdf?se>
- Espín, J. V. (2002). El análisis de contenido: una técnica para explorar y sistematizar información. *XXI, Revista de Educación*, 4, 95-105. Obtenido de <https://rabida.uhu.es/dspace/bitstream/handle/10272/1913/b15141895.pdf?se>
- Hae-Young, K. (2019). Statistical notes for clinical researchers: the independent samples t-test. *Restoraritive Dentistri & Endodontics*, 44(3), e26. doi:<https://doi.org/10.5395/rde.2019.44.e26>
- Lam, R. M. (2016). La redacción de un artículo científico. *Revista Cubana de Hematología, Inmunol y Hemoterapia*, 32(1), 57-69. Obtenido de <https://www.medigraphic.com/pdfs/revcubhemimhem/rch-2016/rch161f.pdf>
- Martínez-Corona, J. I., Palacios-Almón, G. E., & Oliva-Garza, D. B. (2023). Guía para la revisión y el análisis documental: Propuesta desde el enfoque investigativo. *RA XIMHAI*, 19(1), 67-83. Obtenido de <https://raximhai.uaim.edu.mx/index.php/rx/article/view/219/199>
- Matos, Y., & Pasek, E. (2008). La observación, discusión y demostración: técnicas de investigación en el aula. *Laurus*, 14(27), 33-52. Obtenido de <https://www.redalyc.org/pdf/761/76111892003.pdf>
- McHugn, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, 23(2), 143-149. doi:<http://dx.doi.org/10.11613/BM.2013.018>