

La magnitud del efecto como medida de eficacia en intervención educativa

The effect magnitude of as an effectiveness measurement in educational intervention

Sr. Editor:

En la investigación en ciencias de la salud, uno de los principales obstáculos en el proceso de comunicar los resultados de una intervención en el ámbito educativo es la unidad de medida original de las calificaciones pre y posintervención,¹ ya que estos dependen del tipo de cuestionario usado para la evaluación. Por ejemplo, una diferencia de 10 puntos entre el promedio del grupo experimental (GE) y control (GC) después de la intervención puede significar dos cosas diferentes si en un estudio la puntuación máxima del cuestionario es 30 y en otro estudio es 100. Además de ello, también influye la variabilidad de los datos, porque si bien dos grupos pueden tener promedios de magnitud similar, al tener desviaciones estándar (DE) distintas ($M_1: 50, DE_1: 15; M_2: 49, DE_2: 3$) el panorama general cambia.

En el estudio de *Gómez-López* y colaboradores² se reportan los resultados de una intervención para el desarrollo del razonamiento clínico en estudiantes universitarios. Los hallazgos se reportaron en función del *test de significancia estadística de la hipótesis nula* (NHST) y una valoración descriptiva de los promedios de cada grupo. El principal inconveniente con la NHST es que solo brinda información sobre la presencia del fenómeno (p.e., diferencia de medias significativa de estadística o no), pero no acerca de su magnitud.^{3,4} Esto no permite contar con una medida del grado de eficacia mostrada,⁵ lo cual a su vez dificulta la comunicación científica al no presentar una medida estandarizada que pueda ser utilizada en estudios de revisión (p.e., meta-análisis).

Para superar ello, en la actualidad se implementan medidas de magnitud del efecto (ES, por sus siglas en inglés) en los estudios empíricos. En síntesis, la ES se refiere a una medida de la *magnitud* de la presencia del fenómeno, es decir, además de brindar información sobre la *presencia* de una diferencia, indica *qué tan grande* es esta.⁶ Ese indicador permitirá valorar de forma más precisa, por ejemplo, la eficacia de una intervención.

Uno de los procedimientos más usados en el cálculo de la ES para la diferencia entre grupos es la *d* de Cohen.⁷ No obstante, cuando se trata de una intervención en la cual existe GE y GC, se recomienda la delta de Glass (Δ).^{4,5} La Δ consiste en obtener una diferencia estandarizada de los promedios del GE y GC, tomándose

como referencia la DE del GC por ser la más estable. Su expresión matemática es la siguiente:

$$\Delta = \frac{P_{\text{experimental}} - P_{\text{control}}}{DE_{\text{control}}}$$

Donde $P_{\text{experimental}}$, y P_{control} , son los promedios del GE y GC, respectivamente; y DE_{control} es la DE del GC. La valoración de este estadístico depende de su magnitud y se considera .41, 1,15, y 2,7 como ESmínimo recomendable, moderado y fuerte ferguson. Es decir, si la Δ es 1,5 se interpreta como "el promedio del GE supera en 1,5 DE al promedio del GC". Esto es un indicador más preciso de *cuánta diferencia* hay entre los grupos de acuerdo con su distribución (Fig.) ya que una alejamiento de las medias implicaría menos solapamiento entre las distribuciones y por ende mayor diferenciación.

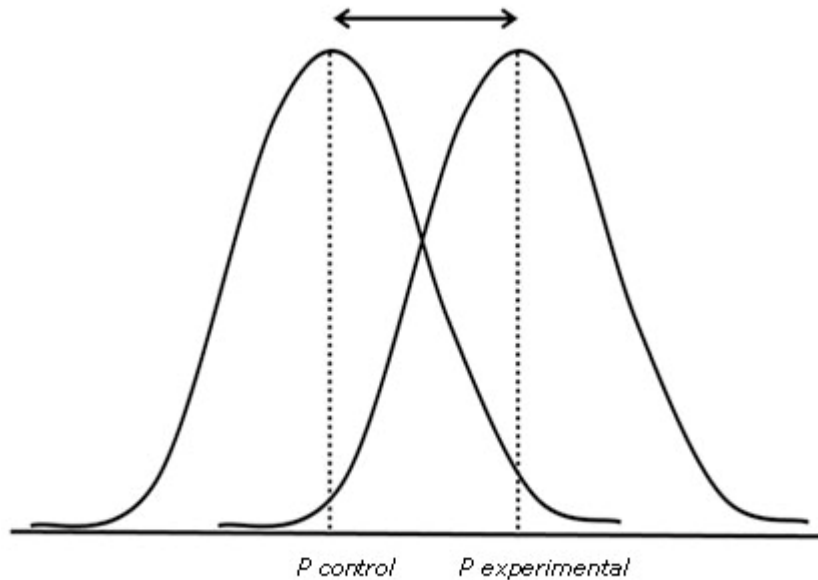


Fig. 1. Diferencia estandarizada de medias.

Para comparar los grupos Gómez-López usaron la prueba t con base en la significancia estadística (rechazar la hipótesis nula si $p < ,05$), pero al disponer de los datos necesarios para el reanálisis (P , DE y n por grupo en la tabla del presente artículo y en la tabla 2 del manuscrito aludido, p. 269), se calculó la Δ para cada momento (pre y pos) a través de un módulo especializado.⁶

Tabla. Comparación entre Grupo experimental y control

	<i>p</i>	DE	n	Δ
Pre-evaluación				
Grupo experimental	63,9	7,2	16	,86
Grupo control	56,4	8,7	22	
Pos-evaluación				
Grupo experimental	97,1	4,4	16	2,28
Grupo control	71,1	11,4	22	
<i>p</i> : Promedio; DE: Desviación estándar; n: tamaño muestral; Δ : Delta de Glass.				

De acuerdo con los nuevos hallazgos, la ES observada en el primer momento (pre-evaluación) podría ser catalogada próxima al nivel *mínimo recomendable*, pero en el segundo momento fue cercana al nivel *fuerte*. Esto concuerda con los resultados presentados en el manuscrito,² donde se indica que la intervención fue efectiva, pero la medida adicional de ES le brinda un mayor sustento empírico al basarse en un indicador estandarizado que puede ser interpretado de acuerdo con criterios establecidos.

En conclusión, incorporar medidas de ES en el proceso de análisis de datos en estudios empíricos es necesario considerando las limitaciones inherentes al NHST, ya que en casos en los cuales se cuente con muestras de tamaño reducido, es más proclive a aceptar la hipótesis nula, y con muestras grandes, a rechazarla. Del mismo modo, resulta útil en el ámbito aplicado y de investigación en ciencias de la salud debido a las diversas aplicaciones que posee,¹ y sobre todo porque comunica en un mismo lenguaje los hallazgos obtenidos.

REFERENCIAS BIBLIOGRÁFICAS

1. Grissom RJ, Kim JJ. Effect size for research. A broad practical approach. New Jersey: Erlbaum; 2005.
2. Gómez-López VM, Rosales-Gracia S, Medellín-del Angel LG, Azcona-Arteaga FJ. Utilidad de una estrategia educativa activo-participativa en el desarrollo del razonamiento clínico en pregrado. Educ Med Super. 2015;29(2):264-72.
3. Coe R, Merino C. Magnitud del Efecto: Una guía para investigadores y usuarios. Revista de Psicología-PUCP. 2002;21(1):47-77.
4. Ferguson CJ. An effect size primer: a guide for clinicians and researchers. Prof Psychol Res Pract. 2009;40(5):532-8.

5. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen.* 2012;141(1):2-18.
6. Dominguez S. Tamaño del efecto: Cálculo mediante un módulo en Visual Basic. *Rev Invest Psicol.* 2013;16(2):235-40.
7. Cohen J. A power primer. *Psychol Bull.* 1992;112(1):155-9.

Recibido: 22 de febrero de 2016.

Aprobado: 1 de julio de 2016.

Sergio Alexis Domínguez-Lara. Instituto de Investigación de Psicología. Universidad de San Martín de Porres. Lima, Perú.
Correo electrónico: sdominguezmpcs@gmail.com