



## Editorial

# Importancia del número de ítems para el cálculo de la consistencia interna

## Importance of the number of items for computing internal consistency

Adalberto Campo-Arias<sup>1</sup> \* Heidi Celina Oviedo<sup>2</sup>

Durante las últimas décadas, el uso de escalas de medición en la investigación y la práctica clínica en salud se ha extendido para la cuantificar de atributos, características, constructos o hacer una aproximación a diagnósticos clínicos.

El empleo de estas escalas, por lo general, es precedido por un complejo y largo proceso de estimación de indicadores de validez y confiabilidad.<sup>1</sup> Una escala válida y confiable debe ser práctica: fácil y breve aplicación, calificación e interpretación.<sup>2</sup> El número de ítems que hacen parte una escala de medición en salud es relevante para el cálculo de algunos indicadores que condicionan la utilidad práctica de un instrumento.<sup>3</sup>

La consistencia interna (CI) es uno de los indicadores de directos de confiabilidad e indirectos de validez más informado para las escalas de medición. La CI mide la homogeneidad de una escala.<sup>4,5</sup> Se aproxima a la correlación entre los ítems que componen la escala, y el error de medición disminuye cuadráticamente a medida que aumenta la CI, por lo que cuanto mayor es, menor es la proporción de error de escala.<sup>4</sup> El cálculo se hace bajo el supuesto que la escala es unidimensional. En este punto es crucial precisar que se debe computar separadamente el indicador de consistencia interna para escala con más de una dimensión. El coeficiente más reconocido es el alfa de Cronbach y está indicado para escalas con patrón de respuesta politómico.<sup>4</sup> El coeficiente de la fórmula 20 de Kuder-Richardson es preferible para escalas con respuesta dicotómica.<sup>6</sup>

La CI es una estimación razonable de la correlación entre los ítems cuando la dimensión tiene un mínimo de dos o tres ítem y un máximo de 15 ítems para los más los expertos más parsimoniosos y de 20, para los más osados.<sup>7</sup> Sin embargo, la CI de dos ítems por el alfa de Cronbach puede subestimarse; aunque, los dos ítems tengan una alta correlación. En estos casos, es mejor utilizar el coeficiente de Spearman-Brown. De tal suerte que la CI por alfa de Cronbach es más predecible para tres ítems.<sup>8</sup>

En condiciones ideales, valores de CI entre 0,70 y 0,95 son bienvenidos.<sup>2,3</sup> Coeficientes inferiores a 0,70 sugieren una baja correlación entre los ítems de la escala y, en consecuencia, posiblemente se relaciona con distintos rasgos latentes.<sup>4</sup> Aunque, una CI alrededor de 0,60 pueden ser aceptables, si se le calcula el intervalo de confianza del 95% por el método de Fisher y el límite superior alcanza el valor de 0,70.<sup>9</sup>

La CI superior a 0,95 es, igualmente, problemática. Estos valores pueden indicar redundancia de ítems por cantidad o calidad. La cantidad de ítems, como se anotó, indica que se calculó la CI a un paquete de 15/20 ítems o más. En resumen, el alfa de Cronbach es altamente sensible al número de ítem.<sup>7,8</sup> Y la calidad de los ítems, evidencia que varios ítems presentan correlaciones superiores a 0,80 entre ellos; correlaciones cercanas a 1,0 significaría sencillamente que ellos miden exactamente el mismo

1. Universidad del Magdalena. Santa Marta, Colombia. Correo: [acamboa@unimagdalena.edu.co](mailto:acamboa@unimagdalena.edu.co) - <https://orcid.org/0000-0003-2201-7404>

2. Universidad Autónoma de Bucaramanga. Bucaramanga, Colombia. Correo: [hoviedo@unab.edu.co](mailto:hoviedo@unab.edu.co) - <https://orcid.org/0000-0003-3239-9450>

componente latente y, en consecuencia, es necesario eliminar todos los ítems sobrerrepresentados.<sup>3,7</sup> Otro factor que puede afectar la CI es el número de opciones de respuesta para cada ítem. Anteriormente, se anotó que el cálculo de CI puede variar según el número de posibles respuestas, los ítems con respuestas politómicas tienden a mostrar mayor CI que los ítems con respuesta dicotómica. Esto indica que la CI se puede aumentar según el número de opciones de respuestas, a mayor número de opciones mayor CI.<sup>3</sup> De la misma forma, es recomendable hacer un ajuste en la CI cuando la escala carece de homogeneidad, es decir, el número de opciones de respuesta posible difiere entre los ítems de una misma escala. Por ejemplo, un ítem ofrece tres opciones de respuesta; otro, cuatro; y un tercero, cinco. En estos casos es necesario ajustar el valor de la CI por número de opciones de respuesta y está más indicado calcular un alfa de Cronbach estratificado.<sup>10</sup> Cuatro o cinco opciones de respuesta para cada ítem suelen ser lo razonable para instrumentos en español o inglés.<sup>3</sup>

En conclusión, la CI se debe calcular sólo a escalas unidimensionales con un mínimo de tres y un máximo de 20 ítems. Con dos ítems, la CI es impredecible; y con más de 20 ítems, se sobredimensiona la CI, más aún si los ítems ofrecen más de cinco opciones de respuesta. Es preferible informar siempre el intervalo de confianza del 95%, particularmente, para los valores de CI cercanos a 0,60.

## DECLARACIÓN SOBRE CONFLICTOS DE INTERÉS

Los autores niegan conflicto de interés que declarar.

## REFERENCIAS

1. Lamm KW, Lamm AJ, Edgar D. Scale development and validation: Methodology and recommendations. *J Int Agric Extens Educ.* 2020;27:24-35. <https://doi.org/10.5191/jiaee.2020.27224>
2. DeVellis RF, Thorpe CT. *Scale development: Theory and applications.* 5th edition. Thousand Oaks: Sage publications; 2022.
3. Streiner DL, Norman GR, Cairney J. *Health measurement scales: A practical guide to their development and use.* Oxford: Oxford University Press; 2024.
4. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16:297-334. <https://doi.org/10.1007/BF02310555>
5. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ.* 2011;2:53-5. <https://doi.org/10.5116/ijme.4dfb.8dfd>
6. Kuder GF, Richardson MW. The theory of the estimation of test reliability. *Psychometrika.* 1937;2:151-60. <https://doi.org/10.1007/BF02288391>
7. Keszei AP, Novak M, Streiner DL. Introduction to health measurement scales. *J Psychosom Res.* 2010;68:319-23. <https://doi.org/10.1016/j.jpsychores.2010.01.006>
8. Eisinga R, Grotenhuis MT, Pelzer B. The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *Int J Public Health.* 2013;58:637-42. <https://doi.org/10.1007/s00038-012-0416-3>
9. Roco-Videla A, Flores SV, Olguín-Barraza M, Maureira-Carsalade N. Alpha de Cronbach y su intervalo de confianza [Cronbach's alpha and its confidence interval]. *Nutr Hosp.* 2024;41:270-1.
10. Tan Ş. Misuses of KR-20 and Cronbach's alpha reliability coefficients. *Educ Sci.* 2009;34:101-12.