



Carta al Director

¿CUÁL ES EL TAMAÑO MUESTRAL ADECUADO PARA VALIDAR UN CUESTIONARIO?

Sr. Editor:

Hemos leído la publicación de Aguilar Navarro y cols. (1) donde se presenta la validación de un cuestionario y en la cual podemos destacar el exhaustivo y fundamentado análisis estadístico realizado. Sin embargo, están ausentes los fundamentos que respaldan el tamaño muestral utilizado. Esto se repite en muchas otras publicaciones, como la de Esteban Figueroa y cols. (2), donde tampoco existe un fundamento que respalde el tamaño muestral pero sí un detallado análisis para la validación a partir de los resultados obtenidos de la muestra.

Si bien no existe un único criterio para determinar el tamaño muestral en el proceso de validación de un cuestionario, sí existen una serie de recomendaciones que creemos que es importante tener siempre a la vista.

Si para la validación realizáramos un análisis estadístico de ítems (media, desviación típica, coeficiente de correlación), bastaría con tener un tamaño muestral que fluctuase entre 50 y 100 sujetos para una primera aplicación, pensando que en una segunda, con el instrumento mejorado, deberíamos tener entre 5 y 10 sujetos por ítem, con un mínimo de 300, ya que de esta forma se pueden tener mayores garantías respecto a la validez del instrumento, dadas las fluctuaciones que pueden tener los estadísticos que se utilicen para la validación (3,4).

En el área de la salud es muy común la utilización de procedimientos factoriales como el análisis factorial confirmatorio (estructura interna). Algunos autores proponen que se deben tener entre 5 y 10 participantes por ítem; otros señalan que tener entre 2 y 3 participantes por ítem es suficiente, siempre y cuando el número total no sea inferior a 200 (5,6). Muestras más pequeñas serían aceptables si la metodología contempla replicar la medición usando diferentes grupos, en donde el número de sujetos sea al menos el doble que el número de ítems, con un total no inferior a los 100 participantes por grupo (7). Snook y Gorsuch señalan, como regla general, que contar con 300 participantes permitiría obtener soluciones fiables independientemente del tipo de análisis factorial realizado (8). Ahora bien, Hair y Colds (9) señalan que el tamaño muestral no depende solo del número de ítems sino también de la carga factorial entre el ítem y el factor o constructo (Tabla I).

Tabla I. Tamaños muestrales necesarios asociados a la significancia de las cargas factoriales

Carga factorial	Tamaño muestral*
0,3	350
0,35	250
0,4	200
0,5	150
0,55	120
0,6	85
0,65	70
0,7	60
0,75	50

*Significancia con un nivel de alfa de 0,05, una potencia del 80 % y errores estándares con un supuesto dos veces mayor que los coeficientes convencionales de correlación (9).

Cuando se desea estimar la fiabilidad de un instrumento, se suelen utilizar el coeficiente test-retest, las formas paralelas o la alfa de Cronbach; las recomendaciones para determinar el tamaño muestral en este caso son las mismas que para el análisis de la estructura interna, es decir, tener entre 200 y 300 participantes (10).

Los fundamentos a considerar para determinar el tamaño muestral son variados. No obstante, antes de tomar alguna decisión, debemos primero asegurarnos de que el tamaño que estimemos permita que todos los segmentos de la población que sean de nuestro interés estén correctamente representados, evitando sesgos por exceso o deficiencia. Si esta consideración no se cumple, el instrumento no logrará su finalidad, a pesar de que se nos señale estadísticamente que estamos en presencia de un instrumento fiable.

Ángel Roco Videla^{1,2}, Marisol Hernández Orellana³
y Omar Silva González⁴

¹Facultad de Ciencias de la Salud. Programa de Magíster en Ciencias Químico-Biológicas. Universidad Bernardo O'Higgins. Santiago, Chile. ²Departamento de Ingeniería Civil. Facultad de Ingeniería. Universidad Católica de la Santísima Concepción. Concepción, Chile. ³Dirección de Informática Educativa. Universidad Autónoma de Chile. Santiago, Chile. ⁴Facultad de Medicina Veterinaria y Agronomía. Universidad de las Américas. Santiago, Chile

Conflictos de intereses: los autores declaran no tener conflictos de intereses.

BIBLIOGRAFÍA

1. Aguilar-Navarro M, Muñoz-Guerra J, Plata MM, del Coso J. Validación de una encuesta para determinar la prevalencia en el uso de suplementos en deportistas de élite españoles. *Nutr Hosp* 2018;35(6):1366-71. DOI: 10.20960/nh.1851
2. Esteban-Figuerola P, Jardí C, Canals J, Arija V. Validación de un cuestionario corto de frecuencia de consumo alimentario en niños pequeños. *Nutr Hosp* 2020;37(1):101-13. DOI: 10.20960/nh.02670
3. Osterlind SJ. *Constructing test Items*. London: Kluwer Academic Publishers; 1989.
4. Martínez-Arias R. *Psicometría: Teoría de los test psicológicos y educativos*. Madrid: Síntesis; 1995.
5. Stevens J. *Applied multivariate statistics for the social sciences*. Hillsdale, NJ: Lawrence Erlbaum; 1992.
6. Kline P. *An Easy Guide to Factor Analysis*. Newbury Park: Sage; 1994.
7. Kline P. *A Handbook of test Construction*. New York: Methuen; 1986.
8. Snook SC, Gorsuch RL. Component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin* 1989;106(1):148-54. DOI: 10.1037/0033-2909.106.1.148
9. Hair JF, Anderson RE, Tatham RL, Black WC. *Análisis Multivariante*, 5a Ed. Madrid: Prentice-Hall; 1999.
10. Clark LA, Watson D. Constructing validity: Basic issues in objective scale development. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research*. American Psychological Association; 2016. DOI: 10.1037/14805-012