

CARTA AL EDITOR

Santa Marta, 24 de junio de 2017

Señores Editores
Revista *Biomédica*

Estimados señores:

Biomédica publicó recientemente un par de artículos sobre el desempeño psicométrico de dos escalas para la identificación de síntomas emocionales, lo cual subraya la importancia de la psicometría o 'clinimetría' en la investigación en salud, así como la necesidad de verificar, una y otra vez, la confiabilidad y la validez de los instrumentos de medición, en particular aquellos que miden constructos en el contexto de la salud mental.

En la publicación anticipada de uno de los artículos mencionados, Simancas-Pallares, *et al.*, presentan una compleja y completa evaluación del desempeño del cuestionario general de salud de 12 ítems (CGS-12) utilizado en estudiantes de Odontología mediante avanzadas técnicas estadísticas en análisis factorial confirmatorio de creciente uso en Colombia (1).

Sin embargo, una revisión cuidadosa de los indicadores de bondad de ajuste muestra que ninguno es adecuado, incluso aquellos que incluyeron corrección del sesgo de respuesta y errores relacionados, dado los altos valores de χ^2 al cuadrado de Bentler-Satorra, lo que cuestiona la validez de constructo del CGS-12, así como la invariancia de los datos de los estudiantes que completaron el instrumento, como concluyó Hankins a partir de la observación de su rendimiento en una muestra grande de personas en la población general (2). Este es un problema menor para instrumentos como el CGS-12, que presenta un aceptable desempeño en cuanto a sensibilidad y especificidad, comparado con el mejor criterio de referencia, la entrevista clínica estructurada (3).

Además, los lectores deben tener presentes las limitaciones del análisis factorial confirmatorio, pues con frecuencia los ajustes de las covarianzas responden a un afán estadístico por encontrar lo que se quiere encontrar, lo cual constituye otro sesgo que, de manera implícita, quebranta el principio de parsimonia, muy valorado en este contexto.

Asimismo, en todas las escalas se conoce el gran problema que los ítems redactados en negativo generan en el español (4). Por último, siempre es necesario recordar que el desempeño 'clínico' de las escalas de medición en salud varía según las características de la población y que no responde a propiedades intrínsecas del instrumento, y que la utilidad de una escala también considera la fácil calificación e interpretación de las puntuaciones (5).

Adalberto Campo-Arias, MD, MSc
Profesor auxiliar, Programa de Medicina, Facultad de Ciencias de la Salud, Universidad del Magdalena, Santa Marta, Colombia
acampo@unimagdalena.edu.co

Referencias

1. **Simancas-Pallares M, Arrieta KM, Arévalo LL.** Validez de constructo y consistencia interna de tres estructuras factoriales y dos sistemas de puntuación del cuestionario general de salud de 12 ítems. *Biomédica*. 2017;37:308-314. <https://doi.org/10.7705/biomedica.v34i2.3240>
2. **Hankins M.** The reliability of the twelve-item general health questionnaire (GHQ-12) under realistic assumptions. *BMC Public Health*. 2008;8:355. <https://doi.org/10.1186/1471-2458-8-355>
3. **Lindert J, Bain PA, Kubzansky LD, Stein C.** Well-being measurement and the WHO health policy Health 2010: Systematic review of measurement scales. *Eur J Public Health*. 2015; 25: 731-40. <https://doi.org/10.1093/eurpub/cku193>
4. **Salazar MS.** The dilemma of combining positive and negative items in scales. *Psicothema*. 2015;27:192-200. <https://doi.org/10.7334/psicothema2014.266>
5. **Sánchez R, Echeverry J.** Validación de escalas de medición en salud. *Rev Salud Pública*. 2004;6:302-18.

Cartagena de Indias, 5 de julio de 2017

Señores Editores
Revista *Biomédica*

Estimados señores:

Ante todo, agradecemos al profesor Campo-Arias su seria y profunda crítica a nuestro artículo publicado recientemente en *Biomédica* (1). Tal y como se comenta en el artículo, es imperativo continuar realizando estudios que den cuenta de las propiedades psicométricas de los instrumentos de medición en salud en cada contexto en el que sean utilizados (2).

El profesor Campo-Arias menciona en su carta que ninguno de los indicadores de bondad de ajuste fue adecuado, afirmación que no se ajusta a la verdad, toda vez que si se revisan los criterios propuestos por Hu, *et al.* (3), para la evaluación del ajuste de modelos derivados del análisis factorial confirmatorio mediante el modelado de ecuaciones estructurales –las cuales son las más empleadas y aceptadas en la literatura científica sobre psicometría– lo encontrado en el modelo unidimensional con la corrección del sesgo de respuesta ofrece un adecuado ajuste que se tradujo en la validez del constructo. Por otro lado, los altos valores de χ^2 al cuadrado obedecieron estrictamente al tamaño de la muestra, pues esta prueba de bondad de ajuste es sensible al tamaño de la muestra (483 sujetos, en nuestro caso) (4).

Además, es importante aclarar que: 1) el valor de χ^2 al cuadrado derivado del análisis factorial confirmatorio no se obtuvo mediante la aproximación de Satorra-Bentler, ampliamente utilizada en modelos de simulación estadística, pero no en este caso, y 2) en el presente estudio no se valoró la invariancia factorial de los modelos investigados, como afirma el autor. Dicha hipótesis se pone a prueba cuando se trata de comprobar la equivalencia de la estructura de factores frente a una variable categórica de interés (sexo, grupo etario, etc.).

Por otra parte, si no se hubieran tenido en cuenta las precauciones estadísticas para la aplicación de los índices de modificación mediante la correlación del error de la covarianza, entonces sí se hubiera caído en un “afán estadístico por encontrar lo que se quiere encontrar”, ya que el lector puede interpretar que se obtuvo un buen ajuste inducido por los investigadores al superar la recomendación para la aplicación de los índices de modificación. No obstante, dado que la hipótesis del estudio fue valorar este modelo (que incluye la correlación de la covarianza de los errores de los ítems negativos), esto no constituye un sesgo, pues la hipótesis del estudio justificaba su aplicación. Además, la parsimonia del modelo no depende del ajuste por correlación de la covarianza de los errores, sino del valor de χ^2 al cuadrado obtenido y de los grados de libertad, los que, a su vez, y como se mencionó anteriormente, se derivan de una prueba estadística sensible al tamaño de la muestra (4).

Como autores estamos de acuerdo con el gran problema que generan en el español las escalas que incluyen ítems redactados en forma negativa, sin

embargo, esto es un problema inherente al diseño de la escala. Así las cosas, el haber demostrado el efecto de los ítems negativos sobre la validez del constructo del cuestionario mediante la correlación de los errores de covarianza a los que dan lugar, permite pensar en el desarrollo de futuros estudios en los que pueda modificarse su redacción y, tal vez, obviar los sofisticados procedimientos estadísticos que se emplean para demostrar las adecuadas propiedades psicométricas del instrumento.

Estamos de acuerdo con el profesor Campo-Arias en que el desempeño de la escala varía según las características de la población, lo que constituye una de las justificaciones de nuestro estudio. Asimismo, y como se menciona en el texto del artículo, el CGS-12 es un instrumento de gran utilidad clínica por su fácil aplicación e interpretación de las puntuaciones.

Miguel Simancas-Pallares, OD, MSc
Profesor asistente, Departamento de Investigaciones, Facultad de Odontología, Universidad de Cartagena
msimancasp@unicartagena.edu.co

Katherine Arrieta-Vergara, OD, MSP
Profesora titular, Departamento de Odontología Preventiva y Social, Facultad de Odontología, Universidad de Cartagena
karrietav@unicartagena.edu.co

Luisa Arévalo-Tovar, OD con especialización
Profesora titular. Departamento de Medicina Oral y Cirugía, Facultad de Odontología, Universidad de Cartagena
larevalot@unicartagena.edu.co

Referencias

1. **Simancas-Pallares M, Arrieta KM, Arévalo LL.** Validez de constructo y consistencia interna de tres estructuras factoriales y dos sistemas de puntuación del cuestionario general de salud de 12 ítems. *Biomédica.* 2017;37:308-314. <https://doi.org/10.7705/biomedica.v34i2.3240>
2. **Beaton DE, Bombardier C, Guillemin F, Ferraz MB.** Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* (Phila Pa 1976). 2000;25: 3186-91.
3. **Hu LT, Bentler PM.** Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling.* 1999;6:1-55. <https://doi.org/10.1080/10705519909540118>
4. **Brown TA.** Confirmatory factor analysis for applied research. New York: Guilford Publications; 2015. p. 461.