

Una modificación del coeficiente alfa de Cronbach por errores correlacionados

A modification of Cronbach's alpha coefficient by correlated errors

Sr. Editor:

La presente carta se centra en el impacto de los errores correlacionados en la estimación de la confiabilidad de los puntajes del *Índice de Bienestar Personal* (PWI-7), reportada en Oyanedel et al¹. En este estudio se halló una confiabilidad satisfactoria del puntaje ($\alpha = 0,85$), pero en vista de la presencia de estos errores correlacionados, la confiabilidad está sobrestimada, tal como lo demostraremos más adelante. Los errores correlacionados están reportados en la figura 1 (p. 1148) y son de magnitudes moderadas (0,47, 0,50, 0,53, y 0,56).

El coeficiente α^2 es el estimador de la confiabilidad de los puntajes que se usa con mayor frecuencia, pero muchas veces se emplea sin considerar sus dos supuestos básicos: tau-equivalencia de los ítems y ausencia de errores correlacionados. Aunque generalmente estos requisitos son difíciles de lograr en la práctica³, la violación de cada uno de ellos sesga la estimación de este coeficiente. En cuanto al primero, se refiere a la similitud de las medias (o cargas factoriales) de los

ítems; y el incumplimiento de esta condición trae consigo la subestimación del α , es decir, que se obtendrán coeficientes bajos. Con relación a la presencia de errores correlacionados, ésta produce una sobrestimación del coeficiente³⁻⁵, pues representa la varianza no relacionada con el constructo que se añade a la estimación de la confiabilidad. La violación simultánea de ambos supuestos da como resultado que el α sea considerado como el límite superior de la confiabilidad (*upperbound*)⁶. De acuerdo a la teoría de la medición, uno de sus supuestos es que los errores (o varianza no relacionada con el constructo) entre los ítems son independientes entre sí^{3,6}.

El incumplimiento del supuesto de errores independientes ha sido revisado anteriormente en la literatura especializada, concluyéndose que puede provenir de causas que van más allá del constructo que originalmente se pretende evaluar. Esto incluye, por ejemplo, factores de irrelevantes de contenido⁷, fraseo de los ítems⁵, cercanía de ítems en el protocolo de evaluación⁶, entre otros. Para controlar este impacto, se han desarrollado fórmulas de corrección del coeficiente α cuando sus ítems muestran errores correlacionados⁸⁻¹⁰. La más reciente⁹ proviene de la equivalencia entre el coeficiente α y el coeficiente Spearman-Brown (ecuación 1), y que es el fundamento para derivar una expresión matemática que considere los errores correlacionados (ecuación 2):

$$r_{xx} = \frac{kr_{ij}}{1 + (k-1)r_{ij}} \quad (1)$$

$$\alpha' = \frac{kr_{ij}}{1 + (k-1)\sqrt{r_{ij}^2 + 2r_{e_{ij}}^2}} \quad (2)$$

donde k es el número de ítems, r_{ij} es la correlación inter-ítem promedio, y $r_{e_{ij}}$ es el promedio de la correlación entre errores. En base a dicha aproximación, se hizo un re-análisis considerando los datos de Oyanedel et al.¹ para obtener una estimación menos sesgada del coeficiente α . Si bien no se conoce r_{ij} , se puede obtenerla usando la ecuación 1; teniendo en cuenta que r_{xx} (0,85) y k (7), el resultado es $r_{ij} = 0,4474$. Por su parte, el promedio de errores correlacionados puede calcularse fácilmente ($r_{e_{ij}} = 0,515$). Finalmente, aplicando la ecuación 2, el α' es de 0,511; el intervalo de confianza¹¹ de esta nueva estimación es 0,435 - 0,580 (95% de confianza). La disminución obtenida es de 39,8%, una cantidad que no podría considerarse insustancial por dos motivos: primero, incrementa el error aleatorio alrededor del puntaje individual de los sujetos evaluados por el instrumento. Y en segundo lugar, el estatus del coeficiente cambia desde un nivel que puede describirse como bueno, hasta un nivel que usualmente se describe como pobre¹².

Efectivamente, la nueva magnitud obtenida es inaceptable incluso para propósitos de investigación y menos aún para propósitos clínicos, en que se requiere alta precisión en la toma de decisiones para el diagnóstico y determinación del cambio luego de una intervención terapéutica^{11,12}. Generalmente, un valor de 0,90 es el mínimo requerido para tomar decisiones clínicas en sujetos, y algo menos para decisiones de grupo^{6,11,12}. Como se puede deducir, esto pone en tela de juicio las conclusiones presentadas por los autores, quienes afirmaron que la confiabilidad es adecuada. Más aún, el grado de confiabilidad corregida indica que el grado de error en el puntaje del PWI-7 afectará directamente a la validez del instrumento (menor a 0,70) ya todo análisis con base en esos datos (comparación de grupos, correlaciones, etc.)⁶.

Los hallazgos en este re-análisis llevarían a los autores del estudio a una re-interpretación de la confiabilidad de los puntajes del PWI-7, así como de su dimensionalidad, dado que debe brindarse una explicación para la presencia de errores correlacionados que sea compatible con lo expresado en la literatura psicométrica. Por otro lado, es plausible que deba considerarse algún otro coeficiente alternativo como estimador de la confiabilidad del PWI-7, pues las diferencias en las magnitudes de las cargas factoriales de sus pueden re-

presentar que la tau-equivalencia no se cumple. Otros coeficientes de confiabilidad, como el coeficiente omega o Feldt-Brennan⁶, pueden ser adecuados en lugar del coeficiente α , pues no requieren del cumplimiento de esta característica.

Sergio Alexis Dominguez-Lara¹, César Merino-Soto¹
¹Instituto de Investigación de Psicología,
 Universidad de San Martín de Porres

Referencias

- Oyanedel JC, Vargas S, Mella C, Páez D. Validación del índice de bienestar personal (PWI - 7) en usuarios vulnerables de servicios de salud en Santiago, Chile. *Rev Med Chile* 2015; 143: 1144-51.
- Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 16: 297-334.
- Dunn TJ, Baguley T, Brunsden V. From alpha to omega: A practical solution to the pervasive problema of internal consistency estimation. *Br J Psychol* 2014; 105 (3): 399-412.
- Green SB, Yang Y. Commentary on coefficient alpha: a cautionary tale. *Psychometrika* 2009; 74 (1): 121-35.
- Yang Y, Green SB. A note on Structural Equation Modeling estimates of reliability. *Struct Equat Modeling* 2010; 17 (1): 66-81.
- Meyer JP. *Reliability*. New York: Oxford University Press; 2010.
- Raykov T. Coefficient alpha and composite reliability with interrelated non homogeneous items. *App Psych Meas* 1998; 22 (4): 375-85.
- Komaroff E. Effect of simultaneous violations of esencial tau equivalence and uncorrelated error son coefficient alpha. *App Psych Meas* 1997; 21 (4): 337-48.
- Pascual-Ferrá P, Beatty MJ. Correcting internal consistency estimates inflated by correlated item errors. *Commun Res Rep* 2015; 32 (4): 347-52.
- Rae G. Correcting coefficient alpha for correlated errors: Is α a lower bound of reliability? *App Psych Meas*. 2006; 30 (1): 56-9.
- Dominguez S, Merino C. ¿Por qué es importante reportar los intervalos de confianza del coeficiente alfa de Cronbach? *Rev Latino Cien Social Niñez Juventud* 2015; 13 (2): 1326-8.
- Merino C, Navarro J, García W. Revisión de la consistencia interna del Inventario de Inteligencia Emocional de Bar-On, EQ-I: YV. *Rev Per Psicol Trab Soc* 2014; 3 (1): 141-54.

Correspondencia:

Sergio Dominguez Lara

Instituto de Investigación de Psicología, Universidad de San Martín de Porres.

Av. Tomás Marsano 242 (5to piso), Lima 34-Perú

sdominguezmpcs@gmail.com

Los autores aludidos ofrecieron la siguiente réplica: Cálculo de confiabilidad a través del uso del coeficiente Omega de McDonald

Reliability calculation using the McDonald's Omega coefficient

Estimado Editor:

En su réplica al artículo “Validación del Índice de Bienestar Personal (PWI) en usuarios vulnerables de servicios de salud en Santiago, Chile”¹, Sergio Domínguez y César Merino argumentan que el estadístico alfa de Cronbach no es un medio adecuado para calcular la confiabilidad de una escala en el caso de la existencia de errores correlacionados en el análisis factorial confirmatorio. Quisiéramos señalar que mayoría de las escalas relevantes son multidimensionales y un alfa satisfactorio generalmente coexiste con varios factores de primer orden y uno general de segundo orden. Es por eso que se plantea, que hay que ir más allá del alfa, utilizar intervalos de confianza de este mediante técnicas de *bootstrap* o re muestreo y usar índices como el Omega de McDonald.

De acuerdo a la propia literatura utilizada por los autores², una alternativa al coeficiente alfa de Cronbach es el coeficiente omega de McDonald (ω). Este coeficiente es calculado a través de los resultados obtenidos en análisis factoriales, de acuerdo a la siguiente fórmula³:

$$\omega = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum \Psi_i}$$

donde λ_i representa la carga factorial, y Ψ_i la unicidad del ítem i .

Considerando esta fórmula, se estimó el coeficiente recurriendo al programa Mplus v. 6.12, tal como el modelo original del artículo. El valor obtenido fue 0,782.

Tal como Domínguez y Merino puntualizaron, el omega obtenido fue menor al del alfa reportado en el artículo, pero de todos modos supera los valores convencionalmente utilizados. En ese sentido, es necesario precisar que PWI-7 no es un instrumento clínico, como afirman Domínguez y Merino, sino que su propósito es determinar niveles de satisfacción vital⁴.

En síntesis, la crítica realizada por Domínguez y Merino es acertada al puntualizar las falencias del estadístico alfa de Cronbach, pero el cálculo del coeficiente omega de McDonald revela una adecuada confiabilidad.

Es importante señalar que si bien el coeficiente omega es más robusto y no exige que los errores no estén correlacionados tiene limitaciones: sus supuestos tam-

poco concuerdan con los datos recogidos usualmente en nuestras áreas, depende del tamaño de la muestra y sobre todo es para “iniciados” ya que no existen paquetes estadísticos amistosos que lo entreguen-SPSS, por ejemplo, no calcula el omega.

Es importante además señalar que una fiabilidad demasiado elevada –como la que los autores plantean convencionalmente para el uso clínico– no asegura la sensibilidad, especificidad y validez predictiva incremental del instrumento. Es más, es ampliamente conocido que indicadores mono ítem como la autoevaluación del estado de salud tienen validez predictiva de la sobrevida, controlando los factores de riesgo clásico⁵. De igual manera, indicadores mono ítem de bienestar subjetivo tienen validez predictiva sobre la sobrevida⁶.

Un coeficiente de fiabilidad muy alto es contraproducente desde el punto de vista de la investigación. Indica redundancia y baja validez de constructo –estos son en su mayoría multidimensionales. Por dar un ejemplo sencillo, si para evaluar emociones positivas elegimos solo alegría y descriptores similares de esta emoción tendremos un alfa (u omega) alto. Si incluimos otras emociones positivas morales como orgullo, otras de baja activación como calma, serenidad, de alta activación como euforia, y emociones de trascendencia, como gratitud, respeto e inspiración, bajaremos la fiabilidad, pero cubriremos de mejor forma el contenido del constructo. Focalizarse solo en la fiabilidad y en estimadores más exigentes es limitar innecesariamente la capacidad de entender la complejidad de los fenómenos estudiados. Implica, en cierta medida, mirar la punta del dedo cuando se está apuntando a la luna. Nuestro interés es contribuir a la evaluación de políticas públicas mediante un instrumento que recoja válidamente todas las áreas relevantes de acuerdo a lo que hoy sabemos sobre satisfacción vital. El PWI 7 muestra internacionalmente amplia validez convergente y de constructo, lo que es cierto también en el caso de los usuarios vulnerables de servicios de salud en Santiago de Chile.

**Juan Carlos Oyanedel¹,
Salvador Vargas², Camila Mella³, Darío Páez⁴**
*Universidad Andrés Bello, Facultad de Educación.
Fundación Centro de Estudios Cuantitativos
(Santiago, Chile).*

²*Fundación Centro de Estudios Cuantitativos
(Santiago, Chile). Escuela de Doctorado, Universidad
de Girona (Girona, España).*

²*Department of Social Policy and Intervention,
University of Oxford (Oxford, Reino Unido)*

²*Facultad de Psicología, Departamento de Psicología Social y Metodología de las Ciencias del Comportamiento, Universidad del País Vasco (San Sebastián, España). Facultad de Administración y Economía, Universidad de Santiago (Santiago, Chile).*

Referencias

1. Oyanedel JC, Vargas S, Mella C, Páez D. Validación del índice de bienestar personal (PWI-7) en usuarios vulnerables de servicios de salud en Santiago, Chile. *Rev Med Chile* 2015; 143: 1144-51.
2. Dunn TJ, Baguley T, Brunsden V. From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *Br J Psychol.* 2014; 105(3): 399-412.
3. McDonald RP. *Test Theory. A unified treatment.* Psychology Press; 2013.
4. The International Wellbeing Group. *Personal Wellbeing Index-Adult (PWI-A). Manual 2013 (5ª ed.).* Melbourne, Australia: Australian Centre on Quality of Life, Deakin University; 2013.
5. Benyamini Y, Idler E. Community Studies Reporting Association Between Self-Rated Health and Mortality. *Research on Aging* 1999; 21 (3): 392-401
6. Diener E & Chan M. (2011) Happy People Live Longer: Subjective Well-Being Contributes to Health and Longevity. *Appl. Psychol. Health Wellbeing* 2011; 3: 1-43.

Correspondencia:
Juan Carlos Oyanedel
Bandera 84, Oficina 112. Santiago de Chile.
juan.oyanedel@fcec.cl