

De regreso a la validez basada en el contenido

Back to content-based validity

JOSÉ VENTURA-LEÓN*.

*Universidad Privada del Norte, Lima, Perú.

En la actualidad es frecuente el desarrollo y validación de un instrumento de medida autoinformado como forma de cuantificar comportamientos adictivos. No obstante, existe una hegemonía de los métodos de ecuaciones estructurales y por ende de la evidencia basada en la estructura interna, en artículos relacionados con las adicciones (Mezquita, Camacho, Suso-Ribera, Ortet y Ibáñez, 2018; Mezquita, Ruiz-Valero, Martínez-Gómez, Ibáñez y Ortet-Fabregat, 2019). A pesar de que, se sabe que existen otras fuentes de evidencia como consecuencias del test, relación con otras variables, procesos de respuesta y contenido, que son considerados un rigor metodológico al momento de investigar adicciones (Fonseca, 2017). En ese sentido, el propósito de esta carta al editor es examinar la importancia de la validez basada en el contenido en el desarrollo y/o adaptación de instrumentos de medida en adicciones.

Supóngase que se desea construir una escala de adicción al amor y un ítem refiere: “*experimento angustia cuando mi pareja se aleja de mí*”, cuya carga factorial es mayor a ,30; pese a ello, la pregunta podría no capturar un comportamiento representativo del universo de conductas de dicho constructo (Cohen y Swerdik, 2001). Por ende, el investigador requiere el juicio y valoración de un experto en la materia, que puede ser considerado como tal, por su amplia trayectoria o por ser una persona reconocida en estos menesteres (Escobar-Pérez y Cuervo-Martínez, 2008).

El procedimiento mediante el cual se revisa la lógica o analiza cuan representativos y relevantes son los contenidos de la prueba para interpretar los puntajes del test, se

denomina *validez basada en el contenido* (American Psychological Association [APA], American Educational Research Association [AERA], y National Council on Measurement in Education [NCME], 2014). De ese modo, revisar la representatividad o relevancia del constructo, puede evitar que exista covariación de información teórica errónea (Haynes, Richard y Kubany, 1995) evitando de esa forma varianza irrelevante asociada al constructo (APA, AERA y NCME, 2014), situación que resulta importante porque un instrumento de medida no solo debe recaer en modelos factoriales, sino en argumentación teórica que sustente si un ítem es representativo o no de un dominio (Bonifay, Lane y Reise, 2017).

En base a lo previamente expuesto, el autor de esta carta ofrecen a los lectores un formato de jueces expertos que puede ser solicitado en su totalidad, sin costo alguno (véase un extracto en Apéndice A), el mismo que está basado en las directrices del APA, AERA y NCME (2014) considerando la relevancia (grado en que el ítem es importante y debe ser incluido en la medición del constructo), representatividad (grado en que el ítem representa el constructo que se pretende medir); además, se incorpora un criterio de claridad (grado en que el ítem es claro y comprensible).

Una vez obtenida las respuestas de los jueces, estas pueden ser cuantificadas mediante la *V* de Aiken, un coeficiente sencillo de calcular y fácil interpretar, tal como se expresa a continuación (Penfield y Giacobbi, 2004):

$$V = \frac{\bar{X} - l}{k}$$

Recibido: Julio 2018; Aceptado: Noviembre 2018.

Enviar correspondencia a:

Jose Ventura-León. Av. Tingo María 1122, Breña, Lima, Telf. (01)6044700 anexo: 3462.

E-mail: jose.ventura@upn.pe

Donde \bar{X} es el promedio de las calificaciones de los jueces; l es la puntuación más baja que es posible obtener y k , es la diferencia entre el mayor y menor puntaje de la escala valorativa. Valores de V cercanos a 1 indican un perfecto acuerdo entre los jueces. Un punto de corte mínimo requerido es ,70 (Napitupulu, Syafrullah, Rahim, Amar y Sucahyo, 2018). Asimismo, en la actualidad se puede establecer intervalos de confianza (IC) para la V Aiken, su expresión matemática se presenta a continuación (Penfield y Giacobbi, 2004):

$$L = \frac{2nkV + z^2 - z\sqrt{4nkV(1-V) + z^2}}{2(nk + z^2)}$$

$$U = \frac{2nkV + z^2 + z\sqrt{4nkV(1-V) + z^2}}{2(nk + z^2)}$$

Donde L es el límite inferior y U , el límite superior; n es la cantidad de jueces; k es la diferencia entre el mayor y menor puntaje de la escala valorativa; V es el valor de la V de Aiken; y z es la distribución estándar elegida, así, 90%, 95% y 99% de confianza le corresponde 1,65, 1,96 y 2,58 respectivamente.

Para la interpretación de los IC, es recomendable que el límite inferior presente un valor $\geq ,70$ (Charter, 2003); a pesar de que se sabe que la amplitud IC depende en gran medida de incremento del tamaño muestral (Penfield y Giacobbi, 2004). Si se desea calcular la V Aiken con sus respectivos IC, se puede solicitar al autor de esta carta una hoja de cálculo en Excel® sin costo alguno o utilizar los siguientes códigos en el programa estadístico R:

```
#####V AIKEN#####
x = 1,90 # media aritmética de las respuestas de los jueces
l = 0 # valor más bajo
s = 3 # valor más alto
k = s-l # rango
v = (x-l)/k ##ecuación v Aiken
v
#####INTERVALOS DE CONFIANZA#####
Z = 1,96 # valor de z al 95%
N = 10 # número de jueces
IC1 = (((2*v*N*K)+(Z^2)))
IC2 = Z*(sqrt((4*N*K*v)*(1-v)+(Z^2)))
IC3 = 2*((N*K)+(Z^2))
INFIC = (IC1- IC2)/IC3
SUPIC = (IC1+IC2)/IC3
INFIC
SUPIC
```

En conclusión, resulta relevante la incorporación de la evidencia basada en el contenido en los instrumentos autoinformados de adicciones, por dos razones: (a) la revisión del contenido de los ítems por jueces expertos previo a la realización de los análisis estadísticos permitirá reducir la varianza irrelevante del constructo; (b) referir que un test es válido solo porque cuenta con evidencia de validez basada en la estructura interna, resulta insuficiente, se requiere de la exploración de más fuentes de validez, una de ellas la del contenido. Finalmente, los códigos en R y el formato de jueces expertos, esperan ser una forma de regresar a la revisión del contenido de los ítems y su coherencia con los postulados teóricos; para de esa forma brindar escalas, cuestionarios o test mejor calibrados en las investigaciones de adicciones.

Financiación

Universidad Privada del Norte.

Conflicto de intereses

Ninguno.

Referencias

- American Educational Research Association, American Psychological Association, y National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bonifay, W., Lane, S. P. y Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science*, 5, 184-186. doi:10.1177/2167702616657069.
- Charter, R. A. (2003). A breakdown of reliability coefficients by test type and reliability method, and the clinical implications of low reliability. *Journal of General Psychology*, 130, 290-304. doi:10.1080/00221300309601160.
- Cohen, R. y Swerdlik, M. (2001). *Pruebas y evaluación psicológicas: Introducción a las pruebas y a la medición*. (4ª ed.). México: Mc Graw Hill.
- Escobar-Pérez, J. y Cuervo-Martínez, A. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en Medición*, 6, 27-36.
- Fonseca, E. (2017). Methodological rigour in the study of addictions. *Adicciones*, 29, 147-149. doi:10.20882/adicciones.994.
- Haynes, S. N., Richard, D. y Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7, 238-247. doi:10.1037/1040-3590.7.3.238.
- Mezquita, L., Camacho, L., Suso-Ribera, C., Ortet, G. y Ibáñez, M. I. (2018). Desarrollo y validación de la ver-

sión corta del cuestionario sobre expectativas de los efectos del alcohol (EQ-SF). *Adicciones*, 30, 271-281. doi:10.20882/adicciones.920.

Mezquita, L., Ruiz-Valero, L., Martínez-Gómez, N., Ibáñez, M. I. y Ortet Fabregat, G. (2019). Desarrollo y validación de la versión breve del cuestionario de motivos de consumo de marihuana (MMM SF). *Adicciones*, 31, 106-116. doi:10.20882/adicciones.979.

Napitupulu, D., Syafrullah, M., Rahim, R., Amar, A. y Sucahyo, Y. G. (2018). Content validity of critical success

factors for e-Government implementation in Indonesia. In *IOP Conference Series: Materials Science and Engineering*, 352, 1-10. doi:10.1088/1757-899X/352/1/012058.

Penfield, R. D. y Giacobbi, Jr, P. R. (2004). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science*, 8, 213-225. doi:10.1207/s15327841mpee0804_3.

Apéndice A

Tabla 1. Parte del formato para jueces expertos.

		Relevancia				Representatividad				Claridad				Sugerencias			
Autoestima positiva*																	
Percepción positiva de sí mismo, observando sus cualidades																	
Nº Ítems																	
1	Me siento una persona tan valiosa como las otras	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
3	Creo que tengo algunas cualidades buenas	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
4	Soy capaz de hacer las cosas tan bien como los demás	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
6	Tengo una actitud positiva hacia mí mismo	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
7	Casi siempre me siento bien conmigo mismo	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
Autoestima Negativa																	
Percepción negativa de sí mismo, tiende a observar aspectos desagradables de su persona.																	
Nº Ítems																	
2	Casi siempre pienso que soy un fracaso	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
5	Pienso que no tengo mucho de lo que estar orgulloso	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
8	Me gustaría tener más respeto por mí mismo	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
9	Realmente me siento inútil en algunas ocasiones	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
10	A veces pienso que no sirvo para nada	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3

Nota: Las alternativas de respuesta en la escala van del 1 al 4 y tiene las siguientes expresiones: (0)Muy en desacuerdo; (1)Desacuerdo; (2)De acuerdo; (3)Muy en desacuerdo; *: Extracto del formato de jueces experto basado en la escala de Rosenberg en español por Atienza, Moreno y Balaguer, 2000.

Firma del Juez Experto

.....

