

Un instrumento de medición y diferentes grupos: ¿cuándo podemos hacer comparaciones válidas?

One measurement instrument and different groups: When can we make valid comparisons?

Señor Editor,

Los procesos de construcción y adaptación cultural de instrumentos son cada vez más comunes en salud. El objetivo de estos procesos es lograr herramientas que sean capaces de medir, de manera válida, ciertos constructos no observados como por ejemplo, calidad de vida, percepción de bienestar o satisfacción. Para esto se llevan cabo rigurosos procesos de traducción, aculturación, análisis de confiabilidad y estructura interna. Sin embargo, ninguno de estos garantiza que los resultados obtenidos sean comparables a través de diferentes grupos (por ejemplo, hombres/mujeres, mayor/menor nivel educacional, mayor/menor edad, diferentes países). Aspectos como la forma en que los grupos conceptualizan la variable que se pretende medir, comprenden los ítems o interpretan la escala de respuesta pueden experimentar variaciones¹. Esta situación limita las inferencias realizadas a partir de los instrumentos de medición.

La comparabilidad o invarianza es una propiedad deseable en todo instrumento de medición y su ausencia indica sesgo. Se dice que existe sesgo en un instrumento cuando los puntajes de los examinados no solo dependen del nivel del rasgo de interés que poseen, sino también del grupo al que pertenecen¹. Diferentes examinados pueden tener diferentes puntajes, pero estas variaciones deben relacionarse principalmente con la variación en el nivel de la variable que se desea medir (por ejemplo, diferentes niveles de calidad de vida) y no del grupo al que pertenecen. Por ejemplo, hombres y mujeres con un mismo nivel en la variable de interés deben poseer la misma probabilidad de alcanzar un determinado puntaje.

La aproximación más utilizada para el estudio de invarianza de instrumentos de medición es la invarianza factorial. Esta aproximación utiliza como base el análisis factorial confirmatorio (AFC) para múltiples grupos¹. El AFC es un análisis multivariado, que pretende explicar las relaciones entre variables observadas (preguntas de instrumento) a través de un número menor de variables no observadas (factores). El AFC se utiliza frecuentemente considerando un único grupo de examinados, pero puede ser extendido a múltiples grupos en forma simultánea, como podrían ser hombres y mujeres o versiones en diferentes idiomas de una escala. Esto permite comparar el comportamiento de un conjunto de preguntas en varios grupos de manera simultánea.

Las pruebas de invarianza consisten en ajustar una

serie de modelos con restricciones sucesivas de algunos de sus parámetros (modelos anidados). Estos modelos son comparados entre sí en base a índices de ajustes¹. Se reconocen cuatro niveles de invarianza: configural, métrico, escalar y estricto (Figura 1). En cada nivel de invarianza se establecen restricciones a ciertos parámetros de interés. Por ejemplo, en el caso de la invarianza métrica se realiza una restricción de equivalencia sobre las cargas factoriales a través de los grupos. Cada modelo se compara con el inmediatamente anterior. Aceptar un nuevo nivel de invarianza implica que todos los anteriores han sido aceptados también. En general para hacer comparaciones entre los grupos a partir de los puntajes obtenidos es necesario contar con un nivel de invarianza escalar.

Para ilustrar esta propiedad consideremos el caso del *Patient Assessment of Chronic Illness* (PACIC). Este instrumento fue desarrollado para medir desde la perspectiva del paciente, la implementación del modelo de Cuidados Crónicos². Pese a que las propiedades psicométricas de PACIC han sido evaluadas en diferentes países del mundo incluido Chile^{2,3}, no ha sido objeto de análisis de invarianza. La evidencia ha mostrado que su estructura factorial varía cuando se aplica a diferentes poblaciones alrededor del mundo. Por ejemplo, algunas versiones han encontrado los cinco factores propuestos originalmente, mientras que otros muestran estructuras con solo un factor². En la versión chilena³, se describen cinco factores, pero la configuración de los ítems asociados a ellos difiere tanto de la propuesta original como de una versión en español previa. Incluso algunos estudios han mostrado que un número menor de ítems sería apropiado. Respecto de las propiedades de la escala de respuesta, una revisión sistemática reciente encontró que la distribución de las repuestas a través de las opciones disponibles variaba en los estudios analizados y que esto podría ser explicado por factores culturales². Estos hallazgos sugieren la ausencia de incluso los niveles más elementales de invarianza de PACIC entre países, lo cual indicaría que la interpretación del constructo varía en diferentes entornos culturales. Por lo tanto, comparar los puntajes obtenidos en Chile con los de otras partes del mundo podría llevar a conclusiones equivocadas.

Alcanzar invarianza no es una tarea sencilla, incluso para sistemas establecidos de medición a nivel internacional⁴. La escasez de invarianza puede deberse al menos a tres fuentes de sesgo: constructo, ítems y método^{4,5}. El sesgo de constructo ocurre cuando los grupos interpretan tales constructos de diferentes formas o cuando la cantidad de ítems es insuficiente para medir el constructo. El sesgo de ítems se relaciona con problemas en la traducción, el contenido de las preguntas o diferencias en como los grupos interpretan ciertas expresiones. El sesgo en el método alude tanto a diferencias en las formas de administración, así como

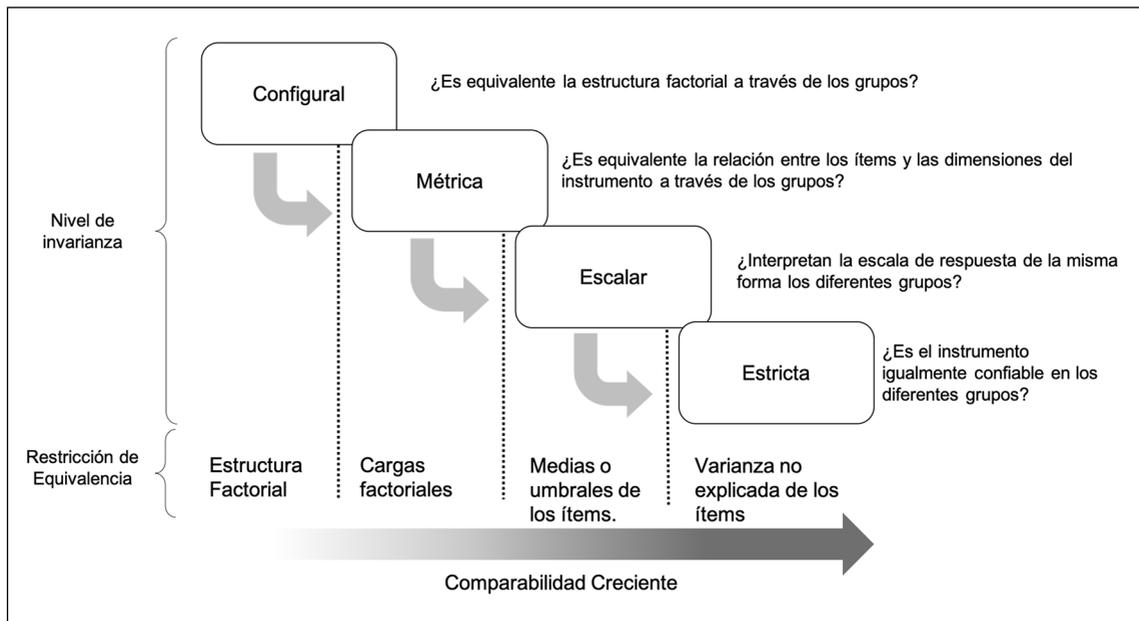


Figura 1. Niveles de invarianza.

también a variaciones en como los grupos interpretan la escala de respuesta. Controlar estas diferentes fuentes de sesgo no siempre resulta tarea sencilla, sin embargo, es necesario comprender que la invarianza no sigue una regla de todo o nada. Es decir, podrían existir partes de la escala que se mantienen invariantes entre los grupos y otras que no. De la misma forma podría no existir invarianza al considerar todos los grupos de interés (por ejemplo, países, diferentes profesiones, rangos etarios) pero sí podría mantenerse entre algunos de ellos.

Los estudios que involucran medición de invarianza en el contexto de la medición en salud aún son escasos. Sin embargo, debido a la relevancia creciente que han adquirido las escalas de auto reporte de pacientes (patient reported outcomes), este tipo de técnicas se ha extendido progresivamente⁶. Esto se ha visto facilitado por su implementación en algunos softwares que hacen su uso más alcanzable, lo cual abre la posibilidad tanto a nuevos estudios transculturales como a posibles mejoras a los instrumentos de medición utilizados comúnmente.

Victor Pedrero¹, Jorge Manzi²

¹Mg. Medición y Evaluación de Programas Educativos. Profesor Investigador Facultad de Enfermería Universidad Andrés Bello.

²PhD Psicología. Profesor Titular Escuela de Psicología Pontificia Universidad Católica de Chile. Director del Centro de Medición de la Pontificia Universidad Católica de Chile.

Referencias

1. Millsap R. *Statistical Approaches to Measurement Invariance*. 1 edition ed. New York: Routledge; 2011 2011/04/14. 367 p.
2. Arditi C, Iglesias K, Peytremann-Bridevaux I. The use of the Patient Assessment of Chronic Illness Care (PACIC) instrument in diabetes care: a systematic review and meta-analysis. *Int J Qual Health Care* 2018; 30 (10): 743-50.
3. Lagos ME, Salazar A, Salas P. Adaptación y validación de la versión chilena del cuestionario "Evaluación de la Atención de Enfermedades Crónicas para pacientes". *Rev Med Chile* 2017; 145 (7): 869-78.
4. Pedrero V, Manzi J. Self-beliefs, engagement and motivation in science and mathematics: Are they universal? *Inter J Educ Res* 2020; 101: 101562.
5. Van de Vijver FJ. Towards a theory of bias and equivalence. In: Harkness J-A, editor. *Cross cultural survey equivalence*. ZUMA-Nachrichten Spezial. Mannheim: ZUMA; 1998. p. 41-65.
6. Edwards MC, Houts CR, Wirth R. Measurement invariance, the lack thereof, and modeling change. *Qual Life Res* 2018; 27 (7): 1735-43.

Correspondencia a:

Victor Pedrero
Facultad de Enfermería, Universidad Andrés Bello. Santiago, Chile.
Victor.pedrero@unab.cl