Colombia Médica

Versión español

English version

Check for updates

CrossMark

EDITORIAL

Adaptación transcultural y validación de escalas e instrumentos de medición en ciencias de la salud

Cross-cultural adaptation and validation of scales and measurement instruments in the health sciences

Andrés Fandiño-Losada

Universidad del Valle, Facultad de Salud, Colombia Médica, Editor en Jefe, Cali, Colombia.

Con este comentario editorial, inicio mi tarea como nuevo Editor en jefe de la revista Colombia Médica de la Universidad del Valle, Cali, Colombia. Agradezco a los comités editoriales y a la Universidad del Valle esta oportunidad de liderar nuestra revista científica de ciencias de la salud. En este editorial quiero referirme a las implicaciones conceptuales y metodológicas del uso de escalas e instrumentos en el área de las ciencias de la salud, ya que es un tema recurrente en los manuscritos que nos envían los autores. De este comentario se excluyen las nosologías, como, por ejemplo, los criterios de Roma IV para el diagnóstico de estreñimiento (intestinal) funcional [1] o los diferentes criterios del Manual Diagnóstico y Estadístico de los Trastornos Mentales, 5.ª edición (DSM-5) [2].

El término "escala" en ciencias de la salud se aplica a dos abordajes diferentes, cada uno con sus propios propósitos y metodologías específicas: 1) las escalas de predicción y 2) las escalas e instrumentos de medición. Las escalas clínicas de predicción se utilizan para estimar, mediante diferentes modelos estadísticos de regresión, la probabilidad de ocurrencia de un evento de salud a partir de un conjunto de variables predictoras. Estos eventos o desenlaces de salud pueden estar ya presentes al observar las variables predictoras en el caso de las escalas de predicción diagnóstica, o bien ser eventos futuros en el caso de las escalas pronósticas[3,4].

Bajo esta conceptualización, no existen dudas razonables sobre la ocurrencia del evento de salud estudiado (por ejemplo, infección intrahospitalaria, enfermedad coronaria, muerte, etc.), ya que estos tipos de eventos cuentan con criterios diagnósticos válidos y métodos diagnósticos basados en estándares de referencia o patrones de oro[5,6]. De esta forma, al contar con patrones de referencia, los métodos diagnósticos siguen procesos de validación mediante pruebas de sensibilidad, especificidad, razones de verosimilitud y de reproducibilidad[7]. Con esta aclaración conceptual y metodológica, a partir de ahora, este comentario editorial se enfocará únicamente en las escalas e instrumentos de medición.

Una variable latente es una característica de la persona o del paciente que no puede observarse ni medirse directamente mediante pruebas físicas, como las fisiológicas, imagenológicas, de laboratorio, etc. La variable latente se corresponde, teórico-conceptualmente, con un constructo teórico también latente, como los estados de salud física (dolor, fatiga, etc.) y mental (deterioro cognitivo, depresión, estrés, etc.), el funcionamiento, la discapacidad, el ambiente psicosocial percibido o la calidad de vida.

Una escala de medición es un instrumento que se utiliza para medir indirectamente una variable latente definida a partir de un constructo teórico-conceptual, mediante las respuestas a un conjunto de ítems, también denominados preguntas o reactivos, que conforman la escala y son observables o medibles directamente, según los fundamentos epistemológicos y metodológicos de la psicometría[8,9]. En las aplicaciones de las ciencias de la salud, las variables observadas, es decir, las respuestas a los ítems de una escala, casi siempre son de naturaleza

ACCESO ABIERTO

Citación: Fandiño-Losada A. **Adaptación transcultural y validación de escalas e instrumentos de medición en ciencias de la salud.** Colomb Med (Cali), 2026; 57(1):e1007448. http://doi.org/10.25100/cm.v57i1.7448

Copyright: © 2026 Universidad del Valle.

Conflicto de interés

Ninguno.

Declaración de financiamiento

La elaboración de este artículo se financió parcialmente con recursos del proyecto de investigación interna CI-11324 de la Universidad del Valle, Cali, Colombia.

Autor de correspondencia

Andrés Fandiño-Losada. Universidad del Valle, Facultad de Salud, Colombia Médica, Editor en Jefe, Cali, Colombia. **Correo electrónico: carlos.fandino@correounivalle.edu.co**

ordinal, como el tipo Likert, donde las respuestas se ubican entre dos polos opuestos, opciones de frecuencia ordenadas que van de "nunca" a "siempre" u otros patrones ordinales, o también de tipo dicotómico "Sí/No".

En la investigación y la práctica clínica, la mayoría de las Mediciones de los Desenlaces Reportados por los Pacientes (MDRP) son escalas que miden variables latentes[10]. Entonces, al no poder contar con estándares de referencia físicos, las escalas y los instrumentos de medición deben seguir un proceso riguroso y sistemático de validación psicométrica, tanto conceptual como empírica, durante su desarrollo original o su adaptación a otro contexto de tipo cultural o lingüístico, con el fin de establecer sus propiedades psicométricas de validez y confiabilidad. La validez se refiere a la capacidad de la escala para medir efectivamente el constructo teórico para el cual fue diseñada, mientras que la confiabilidad se define como la proporción de la variabilidad de los puntajes de una escala que se debe a las diferencias reales entre los sujetos observados en diferentes circunstancias y no a los errores del proceso de medición, que pueden originarse en el instrumento y/o en el observador[11-14].

A manera de ejemplo, revisemos el constructo teórico-conceptual inteligencia y su respectiva variable latente denominada nivel de inteligencia, que no se puede medir mediante una prueba física como un electroencefalograma; por esta razón, se han desarrollado y validado diferentes pruebas de inteligencia, como las *Reynolds Intellectual Assessment Scales* (RIAS) [15], el *Snijders-Oomen Nonverbal Intelligence Test 6-40* (SON-R 6-40)[16], las *Intelligence and Development Scales* (IDS) [17], la *Wechsler Intelligence Scale for Children-Fourth Edition* (WISC-IV) [18] y la *Cattell's Fluid Intelligence Test, Scale 2* (CFT 20-R)[19]. Estas escalas fueron aplicadas en un estudio reciente y sus puntuaciones mostraron una fuerte correlación entre sí y una diferencia mínima entre los promedios de sus puntajes, lo que indica que estas pruebas miden un constructo teórico subyacente muy similar, que se interpreta teóricamente como la inteligencia general[20].

Los resultados del estudio anterior nos indican que los modelos de medición de las escalas e instrumentos no son perfectos, pero su propósito es acercarse lo máximo posible a su respectivo constructo teórico-conceptual [8,9]. La utilidad de estos modelos de medición consiste en que permiten mensurar, con cierto grado de incertidumbre, diferentes constructos latentes sobre el estado de salud o el bienestar de las personas mediante escalas que proporcionan un puntaje numérico que, a su vez, permite clasificarlas en categorías dicotómicas u ordinales, por ejemplo, la severidad de una psicopatología [21].

El marco metodológico de las escalas de medición se fundamenta en la Teoría Clásica del Test y en la Teoría de Respuesta al Ítem [8,9]; en estas teorías, la variable latente es de naturaleza continua. Bajo la Teoría Clásica del Test, se asume que la variable latente es la causa global de las respuestas observadas o medidas en el conjunto de ítems de una escala para todos los individuos. En contraste, bajo el paradigma de la Teoría de Respuesta al Ítem, las características latentes de cada individuo condicionan las respuestas a cada ítem de la escala, lo que permite adaptaciones en su estructura, es decir, en la distribución de los ítems en diferentes dimensiones [8,9].

El proceso de desarrollo, diseño y validación psicométrica original de una escala de medición en un contexto inicial difiere del proceso de traducción, adaptación transcultural, adaptación lingüística y validación psicométrica transcultural de la misma escala en un nuevo contexto, con un idioma, una cultura o ambos diferentes[22]. Por lo tanto, los pasos metodológicos difieren en cada fase de la historia de un instrumento de medición.

Como primer paso, cuando se va a realizar el proceso de adaptación transcultural de una escala o instrumento, se recomienda verificar la coherencia teórico-conceptual del constructo que se pretende medir al utilizar la escala que se está adaptando, así como la validez de contenido de cada uno de sus reactivos. Esto es importante porque en la literatura existen muchas escalas e instrumentos que no se fundamentaron en marcos conceptuales coherentes o que no tuvieron en cuenta las teorías existentes durante su diseño y desarrollo [8,9].

Otro aspecto a considerar durante el desarrollo original o la adaptación transcultural de las escalas de medición es quién informa y determina las respuestas a los reactivos de una escala o instrumento: los Desenlaces Reportados por los Pacientes (DRP) son estados de salud física, mental, funcionamiento, discapacidad y de calidad de vida que los pacientes reportan directamente, sin necesidad de una interpretación de sus respuestas por el personal clínico; asimismo, los instrumentos utilizados para medir estos DRP se denominan MDRP[10]. Las MDRP, al ser respondidas directamente por los pacientes y comunidades, requieren un proceso de adaptación transcultural y lingüística cuando se aplican en un nuevo contexto cultural; por ejemplo, en algunos contextos culturales las prótesis dentales se denominan chapas o cajas; por lo tanto, la adaptación de estos términos es necesaria para lograr la interpretabilidad de las MDRP. Por otro lado, existen instrumentos o escalas que son diligenciados por el personal clínico a partir de su interacción con los pacientes o responsables (por ejemplo, la escala de coma de Glasgow [23], la escala MoCA [24] para evaluar el deterioro cognitivo, etc.); estos deben ser traducidos a otros idiomas, pero no requieren procesos de adaptación transcultural y lingüística para su aplicación en otras poblaciones [25].

Los procesos de traducción, adaptación transcultural y adaptación lingüística están relacionados entre sí, pero conceptualmente no significan lo mismo[21,22]: cuando se realiza una traducción de un instrumento desde su idioma original, por ejemplo, el inglés, a otro nuevo, por decir el español, también se debe realizar el proceso de adaptación transcultural debido al cambio de la cultura y de los estilos de vida en los que se va a aplicar el instrumento traducido. El instrumento puede estar en el mismo idioma, asumamos el español, pero se va a aplicar en una cultura diferente, como de España a Colombia; por lo tanto, requiere una adaptación transcultural y una adaptación lingüística, que incluye adaptar las palabras y modismos locales del mismo idioma al cambiar de un país o región a otro[8,22]. Si se considera que las culturas son similares, por ejemplo, entre Ecuador y Perú, la adaptación transcultural podría no ser necesaria, pero sí podría requerirse la adaptación lingüística. Como ejemplo de aplicación de los conceptos anteriores, en una investigación desarrollada en comunidades afrocolombianas de la Costa Pacífica de Colombia se realizó un proceso de traducción del inglés al español, adaptación lingüística al español de la zona pacífica y adaptación a la cultura afrocolombiana de los siguientes instrumentos: *Hopkins Symptom Checklist* (HSCL-25), *Harvard Trauma Questionnaire* (HTQ) y *PTSD Checklist-Civilian Version* (PCL-C) [26,27].

Después de finalizar los procesos de traducción y adaptación transcultural, las escalas de MDRP y las escalas clínicas deben seguir un proceso riguroso de validación psicométrica, tanto conceptual como empírica, cuando se introducen en nuevas culturas o países, siguiendo las guías de validación psicométrica descritas en revisiones recientes[11-14] o en libros de referencia, como los de Streiner, Norman y Cairney [8] y DeVellis y Thorpe [9]. La validación psicométrica empírica se utiliza para determinar la fiabilidad, la validez y la capacidad de respuesta de una escala en una muestra de sujetos. Las preguntas individuales de los cuestionarios sobre temas de salud orientados a pacientes o comunidades, como las encuestas sobre hábitos saludables y estilos de vida, deben seguir un proceso de traducción y adaptación transcultural y lingüística cuando se van a aplicar en un nuevo contexto cultural; pero no requieren de un proceso de validación psicométrica, a no ser que un conjunto de preguntas, de ese cuestionario, se utilice para conformar una escala de medición.

El estudio de la validez de la estructura factorial de una escala es un componente de la valoración psicométrica empírica, para lo cual se emplean métodos de análisis factorial. Sin embargo, el tipo de análisis factorial que se va a utilizar depende de la fase de desarrollo en la cual se encuentra la escala o instrumento en ese momento: si la escala de medición está en las fases iniciales de su desarrollo, entonces se debe utilizar el Análisis Factorial Exploratorio (AFE), método ateórico que permite identificar la estructura factorial de la nueva escala o instrumento, mediante un abordaje abductivo, a partir de los datos de la muestra de desarrollo y diseño [8,9]. El abordaje abductivo se caracteriza por el uso de analogías para construir descripciones y explicaciones teóricas de la realidad[22].

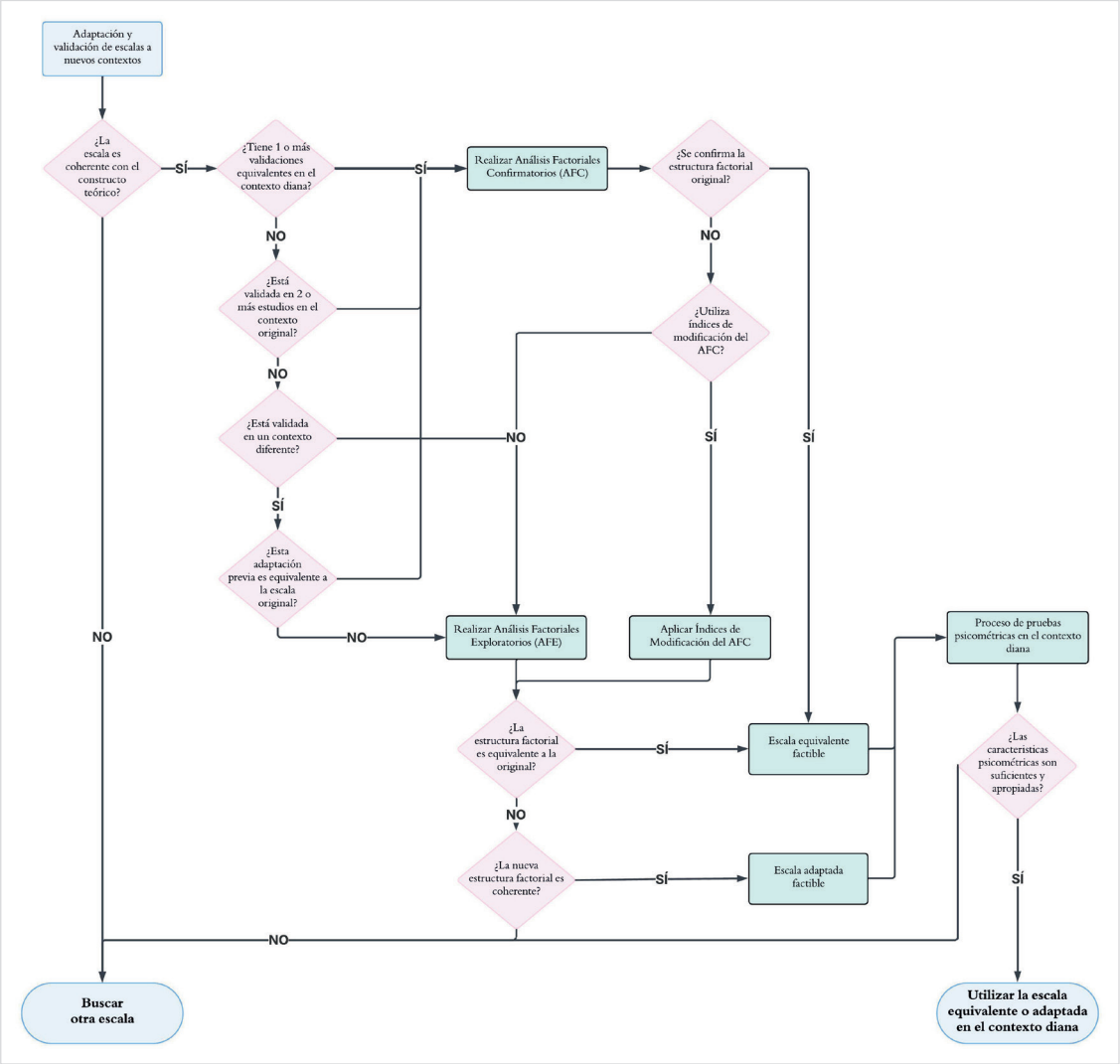

Figura 1. Flujograma para orientar la traducción, la adaptación cultural y la validación transcultural de las escalas de medición de los estados de salud.

Por otro lado, cuando se traduce una escala y se realiza su adaptación transcultural y lingüística, ya existe una teoría previa sobre cómo debe ser su estructura factorial en su contexto original, lo cual implica un idioma y una cultura específicos. Adicionalmente, en un momento histórico posterior, pueden existir otras validaciones psicométricas previas en su contexto original o en contextos diferentes, es decir, en otros idiomas o culturas (Figura 1). Bajo estas circunstancias, el proceso de validación factorial no debería iniciarse con un AFE, sino con un Análisis Factorial Confirmatorio (AFC), un método hipotético-deductivo cuya hipótesis inicial es la estructura factorial original del instrumento que se adapta transculturalmente [8,9]. Si el resultado del AFC no se ajusta a los datos del estudio, se debe realizar un AFE con la muestra de la nueva población para identificar una nueva estructura factorial en este contexto cultural diferente.

No obstante, las nuevas estructuras factoriales sugeridas por el AFE, que incluyen la exclusión de reactivos pertenecientes a los factores o dimensiones originales, deben verificarse posteriormente en otras muestras mediante los AFC [22]. Si el resultado del AFE indica que sólo se debe excluir un ítem, este hallazgo podría deberse a un error estadístico análogo al error tipo II, en el que la carga factorial de ese ítem resultó muy débil en los datos de la muestra del estudio, cuando en realidad el ítem sí pertenece a su factor (o dimensión) teórico. En este caso, se debe considerar conservar la estructura factorial original con todos los ítems que

la componen, para garantizar la comparabilidad de la escala de medición entre diferentes culturas y contextos. Una alternativa metodológica frente al uso del AFE en los procesos de adaptación transcultural de las escalas es la aplicación de los Modelos de Ecuaciones Estructurales Exploratorias (MEEE, ESEM en inglés), que combinan AFC y AFE en un mismo método y aplican los índices de modificación a partir de los AFC[28] (Figura 1).

En las ciencias de la salud, las escalas de respuesta de los ítems son dicotómicas u ordinales, como las respuestas tipo Likert o las opciones ordenadas de frecuencia, y las dimensiones o los factores de un constructo están correlacionados entre sí. Por lo tanto, en los análisis factoriales no se debe utilizar la matriz de correlaciones de Pearson, que es el método por defecto en el software más antiguo, ni las rotaciones ortogonales. En su lugar, se debe utilizar la matriz de correlaciones tetracóricas para las respuestas dicotómicas o, para las respuestas ordinales, la matriz de correlaciones policóricas, y se deben aplicar rotaciones factoriales oblicuas, que permiten la correlación entre los factores de una escala[28,29]. Adicionalmente, como los patrones de respuesta a los reactivos no presentan una distribución normal multivariada, se deben utilizar estimadores robustos, como el estimador de mínimos cuadrados ponderados ajustado por la media y la varianza (WLSMV, por sus siglas en inglés), que está incorporado en el software de análisis factorial más reciente, como JASP©, FACTOR© y los aplicativos lavaan y psych de R©[28,30].

Se recomienda que, en los manuscritos que se vayan a someter a esta y a otras revistas, se incluyan en la sección de Métodos las explicaciones necesarias sobre el desarrollo original, la adaptación transcultural y la validación psicométrica de las escalas utilizadas en cada investigación. Si el manuscrito corresponde a un estudio de adaptación y validación transcultural, se deben seguir las guías de reporte sugeridas por la iniciativa COSMIN versión 2.0[31].

Asimismo, se debe aclarar que el estudio de la validez de constructo se realiza mediante pruebas de hipótesis estadísticas y requiere considerar los subtipos de validez convergente, divergente y discriminante definidos por Streiner *et al.* [8]. En la validez convergente, se espera una correlación alta o moderada entre la nueva escala, por ejemplo, la de dolor, y una escala de un constructo relacionado, como la de calidad de vida. En la validez divergente se espera una correlación baja o nula entre la nueva escala, por ejemplo, el dolor agudo, y una escala de un constructo no relacionado, como la memoria de trabajo. Por último, en la validez discriminante se espera que la nueva escala, por ejemplo, la depresión, tenga puntajes diferenciados entre dos o más grupos de personas, como los pacientes hospitalizados vs. las personas no hospitalizadas[8].

Para orientar el proceso de adaptación y validación transcultural de los instrumentos y escalas de medición de los estados de salud en nuevos contextos, propongo el flujograma presentado en la Figura 1, en el que una versión equivalente significa que la escala, tras la traducción y la adaptación cultural, es teórico-conceptualmente equivalente a la versión original y conserva su estructura factorial. En su conjunto, este flujograma proporciona una ruta sistemática para evaluar la evidencia disponible, la equivalencia estructural y las propiedades psicométricas del instrumento, lo que facilita decidir si utilizarlo directamente, adaptarlo o seleccionar una escala alternativa.

Referencias

1. The Rome Foundation Rome IV criteria. Raleigh: The Rome Foundation. https://theromefoundation.org/rome-iv/rome-iv-criteria/

2. American Psychiatric Association . Diagnostic and statistical manual of mental disorders: DSM-5. 5. Arlington: American Psychiatric Publishing; 2013. [CrossRef] [Google Scholar]

3. Efthimiou O, Seo M, Chalkou K, Debray T, Egger M, Salanti G. Developing clinical prediction models: a step-by-step guide. BMJ. 2024;386:e078276. doi: 10.1136/bmj-2023-078276. [PubMed] [CrossRef] [Google Scholar]

4. van Smeden M, Reitsma JB, Riley RD, Collins GS, Moons KG. Clinical prediction models: diagnosis versus prognosis. J Clin Epidemiol. 2021;132:142–145. doi: 10.1016/j.jclinepi.2021.01.009. [PubMed] [CrossRef] [Google Scholar]

5. Riley RD, van der Windt D, Croft P, Moons KGM, editors. Prognosis research in healthcare: concepts, methods, and impact. Oxford: Oxford University Press; 2019. [CrossRef] [Google Scholar]

6. Ferket BS, van Kempen BJ, Hunink MG, Agarwal I, Kavousi M, Franco OH, et al. Predictive value of updating Framingham risk scores with novel risk markers in the U.S. general population. PLoS One. 2014;9(2):e88312. doi: 10.1371/journal.pone.0088312. [PubMed] [CrossRef] [Google Scholar]

7. Zhou XH, McClish DK, Obuchowski NA. Statistical Methods in Diagnostic Medicine. Hoboken: John Wiley & Sons, Inc; 2009. [Google Scholar]

8. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. 6. Oxford: Oxford University Press; 2024. [CrossRef] [Google Scholar]

9. DeVellis RF, Thorpe CT. Scale development: theory and applications. 5. Thousand Oaks (CA): SAGE Publications; 2022. [Google Scholar]

10. Weldring T, Smith SM. Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs) Health Serv Insights. 2013;6:61–68. doi: 10.4137/HSI.S11093. [PubMed] [CrossRef] [Google Scholar]

11. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health. 2018;6:149. doi: 10.3389/fpubh.2018.00149. [PubMed] [CrossRef] [Google Scholar]

12. Rosellini AJ, Brown TA. Developing and validating clinical questionnaires. Annu Rev Clin Psychol. 2021;17:55–81. doi: 10.1146/annurev-clinpsy-081219-115343. [PubMed] [CrossRef] [Google Scholar]

13. Swan K, Speyer R, Scharitzer M, Farneti D, Brown T, Woisard V, et al. Measuring what matters in healthcare: a practical guide to psychometric principles and instrument development. Front Psychol. 2023;14:1225850. doi: 10.3389/fpsyg.2023.1225850. [PubMed] [CrossRef] [Google Scholar]

14. Stefana A, Damiani S, Granziol U, Provenzani U, Solmi M, Youngstrom EA, et al. Psychological, psychiatric, and behavioral sciences measurement scales: best practice guidelines for their development and validation. Front Psychol. 2025;15:1494261. doi: 10.3389/fpsyg.2024.1494261. [PubMed] [CrossRef] [Google Scholar]

15. Reynolds CR, Kamphaus RW. Reynolds Intellectual Assessment Scales: professional manual. Lutz (FL): Psychological Assessment Resources; 2003. [Google Scholar]

16. Tellegen P, Laros JA. SON-R 6-40: Snijders-Oomen niet-verbale intelligentietest. Göttingen: Hogrefe Verlag; 2011. https://www.hogrefe.com/nl/shop/media/downloads/sample-reports/400139901_InkijkexemplaarHandleiding_Samplepages.pdf [Google Scholar]

17. Grob A, Hagmann-von Arx P. Intelligence and Development Scales-2nd Edition (IDS-2) Oxford: Hogrefe; 2018. [Google Scholar]

18. Wechsler D. Wechsler Intelligence Scale for Children-Fourth Edition. San Antonio: Psychological Corporation; 2003. [CrossRef] [Google Scholar]

19. Cattell RB, Weiß RH. Cattell's Fluid Intelligence Test, Scale 2 (CFT 20-R) Oxford: Hogrefe; 2018. [Google Scholar]

20. Hagmann-von Arx P, Lemola S, Grob A. Does IQ = IQ? Comparability of intelligence test scores in typically developing children. Assessment. 2018;25(6):691–701. doi: 10.1177/1073191116662911. [PubMed] [CrossRef] [Google Scholar]

21. Tsuang MT, Tohen M, Jones PB, editors. Textbook of psychiatric epidemiology. 3. Chichester: Wiley-Blackwell; 2011. [CrossRef] [Google Scholar]

22. Cruchinho P, López-Franco MD, Capelas ML, Almeida S, Bennett PM, Miranda da Silva M, et al. Translation, cross-cultural adaptation, and validation of measurement instruments: a practical guideline for novice researchers. J Multidiscip Healthc. 2024;17:2701–2728. doi: 10.2147/JMDH.S419714. [PubMed] [CrossRef] [Google Scholar]

23. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. Lancet. 1974;2(7872):81–84. doi: 10.1016/s0140-6736(74)91639-0. [PubMed] [CrossRef] [Google Scholar]
24. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc. 2005;53(4):695–699. doi: 10.1111/j.1532-5415.2005.53221.x. [PubMed] [CrossRef] [Google Scholar]

25. Gil L, Ruiz de Sánchez C, Gil F, Romero SJ, Pretelt BF. Validation of the Montreal Cognitive Assessment (MoCA) in Spanish as a screening tool for mild cognitive impairment and mild dementia in patients over 65 years old in Bogotá, Colombia. Int J Geriatr Psychiatry. 2015;30(6):655–662. doi: 10.1002/gps.4199. [PubMed] [CrossRef] [Google Scholar]

26. Santaella-Tenorio J, Bonilla-Escobar FJ, Nieto-Gil L, Fandiño-Losada A, Gutiérrez-Martínez MI, Bass J, et al. Mental Health and psychosocial problems and needs of violence survivors in the Colombian Pacific Coast: a qualitative study in Buenaventura and Quibdó Prehosp Disaster Med. 2018;33(6):567–574. doi: 10.1017/S1049023X18000523. [PubMed] [CrossRef] [Google Scholar]

27. Bonilla-Escobar FJ, Fandiño-Losada A, Martínez-Buitrago DM, Santaella-Tenorio J, Tobón-García D, Muñoz-Morales EJ, et al. A randomized controlled trial of a transdiagnostic cognitive-behavioral intervention for Afro-descendants' survivors of systemic violence in Colombia. PLoS One. 2018;13(12):e0208483. doi: 10.1371/journal.pone.0208483. [PubMed] [CrossRef] [Google Scholar]

28. Lloret-Segura S, Ferreres-Traver A, Hernández-Baeza A, Tomás-Marco I. El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. An. Psicol. 2014;30(3):1151–1169. doi: 10.6018/analesps.30.3.199361. [CrossRef] [Google Scholar]

29. Baglin J. Improving your exploratory factor analysis for ordinal data: a demonstration using FACTOR. Practical Assessment, Research & Evaluation. 2014;19(5):1–15. doi: 10.7275/dsep-4220. [CrossRef] [Google Scholar]

30. Rogers P. Best practices for your confirmatory factor analysis: A JASP and lavaan tutorial. Behav Res Methods. 2024;56(7):6634–6654. doi: 10.3758/s13428-024-02375-7. [PubMed] [CrossRef] [Google Scholar]

31. Gagnier JJ, de Arruda GT, Terwee CB, Mokkink LB. Consensus group COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures: version 2.0. Qual Life Res. 2025;34(7):1901–1911. doi: 10.1007/s11136-025-03950-x. [PubMed] [CrossRef] [Google Scholar]