

Análisis de factores: fundamentos para la evaluación de instrumentos de medición en salud mental*

Adalberto Campo-Arias¹

Edwin Herazo²

Heidi Celina Oviedo³

Resumen

La validación de un instrumento de medición en salud mental es un proceso complejo que habitualmente se inicia con la estimación de la confiabilidad, para posteriormente hacer la aproximación a la validez. El análisis de factores es una forma de conocer el número de dimensiones, dominios o factores de un instrumento de medición, y, por lo general, se relaciona con la validez del constructo de una escala. El análisis puede ser exploratorio o confirmatorio, y ayuda a seleccionar los ítems con mejor desempeño. Para un aceptable análisis de factores es necesario seguir algunos pasos y recomendaciones, hacer pruebas estadísticas y contar con una muestra adecuada de participantes.

Palabras clave: Psicometría, escalas, análisis factorial, estudios de validación, revisión.

Title: Factor Analysis: Principles to Evaluate Measurement Tools for Mental Health

Abstract

The validation of a measurement tool in mental health is a complex process that usually starts by estimating reliability, to later approach its validity. Factor analysis is a way to know the number of dimensions, domains or factors of a measuring tool, generally related to the construct validity of the scale. The analysis could be exploratory or confirmatory, and helps in the selection of the items with better performance. For an acceptable factor analysis, it is necessary to follow some steps and recommendations, conduct some statistical tests, and rely on a proper sample of participants.

Key words: Psychometrics, scales, factor analysis, validation studies, review.

* El Instituto de Investigación del Comportamiento Humano financió esta revisión.

¹ Médico, especialista en Psiquiatría, magíster en Salud Sexual y Reproductiva. Líder del Grupo de Investigación del Comportamiento Humano, Instituto de Investigación del Comportamiento Humano, Bogotá, Colombia.

² Médico, especialista en Psiquiatría, magíster en Bioética. Miembro del Grupo de Investigación del Comportamiento Humano, Instituto de Investigación del Comportamiento Humano, Bogotá, Colombia.

³ Médica, especialista en Psiquiatría. Miembro del Grupo de Investigación del Comportamiento Humano, Instituto de Investigación del Comportamiento Humano, Bogotá, Colombia.

Introducción

Una *escala* es un instrumento de medición compuesta por un grupo de elementos, ítems, incisos, reactivos, preguntas o puntos que exploran una o más dimensiones, dominios o factores de un constructo teórico, como la mayoría de los que se abordan o trabajan en salud mental (1).

La validación o evaluación del desempeño psicométrico de las escalas, en salud mental y en otras áreas del conocimiento o disciplinas, es un proceso continuo, complejo y extenso (2). La validación de un instrumento habitualmente se inicia con la estimación de la confiabilidad, mediante prueba de consistencia interna, estabilidad y concordancia, según lo indique el tipo de instrumento de medición (3). Posteriormente, si se logran valores aceptables en las pruebas anteriores, es necesario estimar la validez (1).

La estimación de la validez es un proceso permanente, aún más complejo que valorar la confiabilidad. La confiabilidad es un paso necesario para iniciar el proceso de estimación de la validez, dado que un instrumento altamente confiable no es necesariamente válido y un instrumento válido carece de utilidad práctica si no es lo suficientemente confiable (4).

No obstante, a las precisiones anteriores en el proceso de validación de una escala, y de estar seguro de la confiabilidad, luego es indispensable conocer la validez (5). Con frecuencia,

este proceso se inicia con la estimación de la consistencia interna de una escala (α de Cronbach o coeficiente de Kuder-Richardson) (6,7), y posteriormente, con los mismos datos se explora la dimensionalidad de esta mediante el análisis de factores (8). La consistencia interna es una propiedad necesaria, pero no suficiente para garantizar unidimensionalidad de una escala, esto quiere decir que la escala se compone de un único factor (9,10).

El análisis de factores es una de las varias formas de estimar la validez de constructo de un instrumento, de mostrar en forma matemática cómo la construcción teórica de un concepto abstracto complejo se refleja en el patrón de respuesta de un grupo poblacional (11-13). En teoría, si un instrumento o cuestionario refleja en forma importante un constructo debe mostrar las dimensiones que propone la conceptualización teórica, y con ello se estima en parte la validez general de la escala (14-16).

Se espera que una escala esté formada por un máximo de 20 ítems (3,9). El análisis de factores es útil en la construcción de una escala o en la redefinición del número de los ítems, es decir, ayuda a decidir qué puntos deben eliminarse o mantenerse en una escala y conservar utilidad práctica (14). Se excluyen los reactivos que muestran coeficientes bajos en la solución factorial (17). La eliminación de estos reactivos con bajos coeficientes mejoran, igualmente, propiedades psicométricas, como la

consistencia interna, que siguen el mismo principio matemático para el cálculo (18,19). Es importante recordar que el análisis de factores, en el proceso de validación de escala y en otros usos en investigación, debe incluir un número no superior a 30 ítems (20).

Inicialmente, se sugería el uso de análisis de factores solo para escalas politómicas (tipo Likert); sin embargo, en la actualidad se considera que los instrumentos con patrón de respuesta dicotómico o binomial, igualmente, pueden ser objeto de análisis de factores con resultados adecuados o aceptables (21).

Es necesario precisar que existen diferentes tipos de validez: aparente, contenido, constructo y criterio (22,23); sin embargo, todas estas formas de validez ayudan, en mayor o menor grado, a conocer la validez práctica del constructo (22). Para otros instrumentos solo es posible la estimación de algunas de estas formas de validez. Así mismo, es bien conocido que para algunos constructos no existe un criterio de referencia (*gold standard*) universalmente aceptado y, en consecuencia, no se puede estimar la validez de criterio que representa el tipo de validez más útil desde una perspectiva clínica (24,25).

El objetivo de esta revisión es presentar a estudiantes y profesionales de la salud mental con conocimientos mínimos de estadísticas los fundamentos teóricos y prácticos para la realización de un análisis de

factores a instrumentos de medición en salud como parte del proceso de validación.

Principios del análisis de factores

El análisis de factores permite explicar un número amplio de ítems que hacen parte de una escala con un número reducido, mucho menor, de nuevas variables (factores), sin pérdida importante de información de los ítems originales (20). A estas pocas nuevas variables se les conoce comúnmente con el nombre de *factores*; igualmente, se les llama *dimensiones* o *dominios* (26).

El análisis de factores parte de la suposición de que existe un componente o parte común, *comunalidad*, en los elementos que hacen parte de una escala. De la misma forma, se entiende que los ítems exploran componentes o rasgos latentes subyacentes, no observables en forma directa (27). Así mismo, se asume que cada punto presenta una parte única variable, no común, *unicidad*, que explora un aspecto particular del constructo, no presente en los otros elementos o ítems. Se considera que las partes únicas de cada ítem son independientes entre sí (20).

Existen dos tipos de análisis de factores. El primero, el análisis de factores exploratorio (AFE) que se realiza para conocer la manera de representar en forma sencilla (parsimonia) cómo se relaciona entre sí un número determinado de ítems, sin conocer cuántos factores pueden

observarse, y sin nombres previstos para dichos factores. Los factores solo se identifican o retienen durante el proceso de análisis (21,28). Y el segundo, el análisis de factores confirmatorio (AFC) se comienza con un número predeterminado de factores, que ya tienen un significado específico, los ítems que hacen parte de cada factor o si los factores se encuentran alta o pobremente correlacionados. Y al final, se busca conocer la bondad del ajuste para el modelo final que se acepte (21,28). Para algunos autores, el AFC es un simple paso más, que realmente no aporta mucho en el proceso de validación de un instrumento si el EFE se llevó a cabo en forma adecuada y cuidadosa (28,29).

Pasos para llevar a cabo un análisis de factores

Para realizar un análisis de factores que garantice la identificación de los factores más importantes de un instrumento es necesario seguir unos pasos mínimos e interpretar en forma adecuada las pruebas estadísticas más usadas (29).

Estimación de la correlación entre los puntos

Para comprobar el grado de correlación entre los puntos que hacen parte de una escala existen diferentes métodos o pruebas estadísticas. Los más conocidos o usados son la prueba de esfericidad de Bartlett (30), la prueba de adecuación de la mues-

tra de Kaiser-Meyer-Olkin (31) y el coeficiente determinante. Estos coeficientes indican si existe una importante correlación entre los ítems y que es probable que estos exploren un aspecto común, es decir, que se agrupen en uno o más dominios o factores, y, en consecuencia, se puede llevar con alguna probabilidad de éxito el análisis de factores. Las pruebas sugieren que es probable encontrar factores, mas no garantizan que la solución factorial deje satisfecho a los investigadores (32). Paquetes estadísticos como SPSS y STATA tienen disponibles estas pruebas (33,34).

La prueba de esfericidad de Bartlett parte de la hipótesis nula de que los ítems que hacen parte de la escala son independientes. Mientras que la hipótesis alterna supone que dichos ítems son dependientes, es decir, muestra una alta correlación entre ellos y pueden agruparse en uno o más factores de un constructo. Para aplicar esta prueba es necesario observar una distribución normal de las puntuaciones en todos los ítems. En forma sencilla, esta prueba da un valor de χ^2 , unos grados de libertad que son iguales al número de puntos que forma la escala multiplicado por el número de variables menos 1 dividido por 2 ($k - 1/2$) y un valor de probabilidad (p). Para rechazar la hipótesis nula se espera un valor alto de χ^2 y un valor de probabilidad menor del 5% (30).

La prueba de adecuación de la muestra de Kaiser-Meyer-Olkin (KMO) es un coeficiente que resulta

de la comparación de las correlaciones que se observan y las correlaciones parciales, es decir, es cuantificación de la correlación existente entre todos los ítems analizados. Si los puntos miden el mismo aspecto deben mostrar altas correlaciones entre sí. El coeficiente KMO puede estar entre 0 y 1. Los valores inferiores a 0,50 se consideran inaceptables; entre 0,50 y 0,59, pobre; entre 0,60 y 0,69, bajo; entre 0,70 y 0,79, modesto; entre 0,80 y 0,89, bueno, y entre 0,90 y 1,0, excelente. Para análisis de factores se prefieren valores superiores a 0,70 (31). Los valores inferiores indican una limitada correlación entre los puntos y, en consecuencia, el número de factores que se obtendrá será casi igual al número de puntos que se estudian. Esto no tiene sentido si parte del hecho de que el objetivo final del análisis de factores es la reducción significativa de un grupo grande de puntos a unos pocos factores (35).

Finalmente, el coeficiente determinante puede oscilar entre 0 y 1. Para alcanzar una solución factorial satisfactoria se espera, por lo general, observar que este se acerque a la unidad. No obstante, coeficientes superiores a 0,60 indican que se puede seguir en el proceso de encontrar por lo menos un factor común en un conjunto de ítems (35).

Métodos de extracción de factores

Existen varios métodos estadísticos para la extracción de los factores

más importantes de un conjunto de puntos. Aquellos que más informan las publicaciones científicas son el método de componentes principales y el método de análisis de factores (28,35).

Método de componentes principales

El método de componentes principales se recomienda para los AFE, dado que se inicia sin una hipótesis previamente establecida del posible número de factores por observar o extraer y de los eventuales elementos o ítems componentes (19). Para algunos autores este no es un verdadero método para identificar la relación existente entre los ítems de una escala, sino una técnica para resumir o reducir en un conjunto en un número menor (28). Sin embargo, se espera que el número de factores sea sustancialmente inferior que el número de puntos iniciales (factores < ítems). Carece de sentido el análisis de factores si el número final de factores es muy cercano al número de variables iniciales. Se entiende que cada uno de los factores es una combinación lineal de todos los reactivos y el 100% de la variabilidad de los elementos es responsable de los factores (20,36).

Método de análisis de factores

Los métodos de análisis de factores agrupan un conjunto de técnicas de máxima verosimilitud (la mejor conocida e informada), mínimos cuadrados

no ponderados, mínimos cuadrados generalizados, factor de eje principal, factor α y factor en imagen. A diferencia del método de componentes principales, estas aproximaciones parten de una hipótesis predefinida acerca del número de factores por considerar en el modelo de análisis y considera que existe interdependencia entre estos (14,28,35). Las referencias para fortalezas y debilidades de estas técnicas son pocas y no confiables (35). Este método toma como base los valores de la comunalidad de cada ítem. En teoría, esta comunalidad se relaciona con la varianza de cada ítem, debido a la influencia que ejercen en el factor, tanto por la unicidad como debida al error aleatorio (36). De tal suerte que estos métodos se indican para la realización de AFC o cuando se pretende redefinir o mejorar la calidad de una escala (21,32,37).

Tipos de rotaciones

La rotación de los factores se usa con el único propósito de lograr una mejor interpretación de los factores que se observan durante el análisis. Las soluciones no rotadas son difíciles de interpretar (28). Mediante la rotación se logran polarizar los coeficientes, se pretende llevarlos hacia los extremos, hacia cero o hacia uno, cuando existe más de un factor. La rotación hace que algunos ítems muestren coeficientes más altos en un factor y más bajos en otros. Esto ayuda a visualizar y

entender mejor la solución factorial, y ver con más claridad en qué factor un reactivo muestra el coeficiente o el peso más alto; finalmente, a decidir a cuál dominio se le puede asignar el aporte que hace. Es importante tener presente que la rotación no induce cambios en las comunalidades y la equivalencia matemática en la solución no rotada (28,32). En pocas palabras, la rotación se realiza para simplificar la interpretación de estructura de factores observada y dar nombre a los factores según los ítems muestren mayores coeficientes (20,28).

Existen dos tipos de rotaciones: ortogonales y oblicuas. Las rotaciones ortogonales parten de la suposición de que algunos elementos que hacen parte de un conjunto tienen una baja correlación entre ellos y, en consecuencia, los factores que se identifiquen mostrarán pobre correlación; es decir, que prácticamente se comportarán como dos constructos distintos. Las rotaciones ortogonales disponibles son la tipo varimax (la más usada) y la equamax (14,26,29, 35).

Por su parte, las rotaciones oblicuas dan por sentado que algunos puntos que hacen parte de un conjunto muestran altas correlaciones y, por consiguiente, los factores que se retengan igualmente mostrarán altas correlaciones entre ellos (20). Las rotaciones oblicuas son la rotación oblimin, la promax (la más popular en este grupo) y la quartimax. Este método es más eficiente para la iden-

tificación de una estructura simple de factores agrupados en una sola escala (20,26).

Criterios para considerar un factor como importante o relevante

El objetivo de análisis de factores es contar con un número significativamente menor de dominios que recojan la esencia de los ítems que se incluyeron inicialmente, de tal suerte que los factores retenidos deben conservar en gran parte la esencia de la medición original. En las mejores condiciones se pretende evitar tanto el exceso (sobrerretención) como el escaso (subretención) número de factores extraídos. Existen varios criterios para considerar cuáles son los factores más importantes o relevantes que se deben retener en una solución de factores (38-41).

Número de ítems por factor

Dada la complejidad de algunos constructos, es difícil abarcar en forma significativa una dimensión con un número reducido de ítems. Así mismo, es necesario tener presente que la reproducibilidad de un factor es inversamente proporcional al número de reactivos que lo componen (20). Existe el consenso de que un factor es realmente importante, y digno de retener, si lo forman por lo menos entre tres y cinco reactivos (20,26,28,36,42). Los dominios con menos de tres ítems muestran poca solidez teórica

y alta inestabilidad (baja reproducibilidad) en distintas mediciones (28).

Valores de los coeficientes

Otro aspecto que se considera al momento de elegir los factores más notorios en una solución de factores es observar los valores que alcanzan los coeficientes de cada punto en cada uno de los factores (43). Algunos autores sugieren tomar en forma arbitraria como punto de corte coeficientes iguales o superiores a 0,300 para estimar que un ítem tiene un peso importante en un factor en particular, con base en el valor que se toma como significativo para las correlaciones bivariadas de Pearson o Spearman (28,35). Otros investigadores sugieren que los coeficientes superiores a 0,71 son excelentes; a 0,63, muy buenos; a 0,55, buenos; a 0,45, aceptables; e inferiores a 0,32, pobres (27). Sin embargo, Stevens propone considerar el tamaño de la muestra al momento de tomar el punto de corte para considerar un coeficiente como significativo. El punto de corte mínimo debe ser igual a $5,152$, dividido entre la raíz cuadrada del tamaño de la muestra menos dos ($5,152/\text{raíz cuadrada de } n-2$) (44).

La fórmula surge del hecho de que la distribución de las correlaciones sigue en forma aproximada la curva de distribución normal gaussiana cuando el número en mediciones es mayor a 100 y 2,576 es el valor

de z para un valor de significación del 1%, que se multiplica por dos, para un valor de 5,152 (43). El uso de uno u otro valor crítico para los coeficientes, por lo general, queda a preferencia o conveniencia de los investigadores (27) y muchas veces guarda relación con el tamaño de la muestra del estudio (28,35).

Valores propios

El *valor propio* es el coeficiente o índice de la varianza de cada uno de los dominios identificados en la solución de factores. Programas estadísticos, por ejemplo SPSS, muestran los valores propios para todos los posibles factores. El número de factores que presenta la solución siempre es igual al número de ítems que se incluyeron en el análisis (20).

Es un contrasentido retener un número de factores igual al número de reactivos iniciales. Para esta retención se considera con mucha frecuencia el valor propio. Los valores propios en una solución de factores pueden estar entre cero y un valor cercano al número de elementos incluidos en el análisis de factores, si se tiene en cuenta que la suma de todos los valores propios es igual al número de elementos que componen la escala (43).

Criterio de Kaiser

La mayoría de los paquetes estadísticos retienen por defecto los factores que muestran valores pro-

prios superiores a 1,0; es el llamado criterio de Kaiser. Un factor con un valor propio inferior a 1,0 induce una varianza menor que un ítem individual (31). El criterio de Kaiser tiene un par de limitaciones. La primera, siempre que se utiliza existe alta probabilidad de sobrestimar el número de factores por retener. Y la segunda, este criterio es muy sensible al número de elementos que se incluyen en el análisis. Así, es posible que se identifique un número excesivo de factores cuando se incluyen 50 o más ítems, o se observe una cantidad reducida de factores cuando se realiza el análisis con 20 o menos reactivos (35,40,41,43).

Dado el principio general del análisis de factores, reducir al máximo posible el número de reactivos, es preferible subestimar que sobrestimar el número de factores retenidos; por lo anterior, se sugiere que se determine en forma bastante razonable el número de factores latentes presentes solo a un conjunto de ítems igual o inferior a 30. Los análisis de factores con más de treinta ítems frecuentemente son difíciles de interpretar en el contexto de la teoría que sostiene la construcción de la escala (20).

Criterio de Gorsuch

Dada las dificultades inherentes a tomar como valor de retención el criterio de Kaiser, en particular la sobrestimación de los factores relevantes, Gorsuch estimó y propuso

que valores propios mayores de 1,41 sí indican de manera significativa los factores que agrupan las características más generales y significativas de un constructo, y, en consecuencia, son los únicos que se deben retener en análisis de factores, como los dominios más importantes o notables del constructo en estudio (26). Con este criterio, es usual la retención de un único factor que condensa el componente central de cualquier constructo (43).

Varianza que explica los factores extraídos

Antes se anotó que el valor propio de cada factor puede estar entre cero y el número de ítems que hacen parte de la escala, y que la suma de todos estos es igual al número de reactivos (43). Por lo tanto, si se multiplica el valor propio por 100 y se divide entre el número de elementos, se conoce el porcentaje de la varianza total que explica dicho factor. Es evidente que la suma de todos los factores posibles da cuenta del 100% de la varianza (35).

Otros autores sugieren que deben retenerse aquellos factores que expliquen por lo menos el 5% de la varianza total (45). Sin embargo, este criterio tiene el gran inconveniente de que la varianza que explica cada factor es directamente proporcional al número de reactivos que se incluyen en el análisis (35). Si se quieren explorar los factores latentes entre un grupo de 20 ítems, teóricamente,

desde el inicio del análisis, cada ítem es responsable al menos del 5% de la varianza total y, por consiguiente, con este criterio se tiende a sobrestimar el número de factores identificados con instrumentos con un número inferior a 20 reactivos (43).

Sin embargo, la varianza total que explica los factores retenidos es de gran importancia para la solución de factores. Se considera que un análisis de factores es aceptable si la suma de las varianzas de los factores que se retienen es igual o superior al 50% (20,40). Con frecuencia, el primer factor es el que muestra el valor propio más alto y, por consiguiente, explica el mayor porcentaje de la varianza de la solución; como se anotó antes, recoge la característica esencial del constructo en estudio (43). No obstante, en algunas situaciones, como al realizar AFE, pueden ser admisibles soluciones factoriales que den cuenta de al menos el 30% de la varianza total (46).

Gráfica de los valores propios de los factores

El criterio o prueba de la pendiente de Cattell (*scree plot*) para la determinación de los factores por retener se basa en la gráfica de los valores propios que muestra cada factor. Se forma una curva en la que aparece una inflexión en el punto en que la pendiente tiende a formar una línea horizontal. Esto ocurre por lo general cuando los valores propios se aproximan a cero. El último factor

por retener es el que se encuentra antes de la inflexión. No obstante, la determinación de la inflexión se hace por simple observación, lo que le da un toque algo subjetivo a esta prueba (47). Los paquetes estadísticos, verbigracia SPSS, presentan la gráfica si se solicita (34).

Sin duda, la mejor solución de factores posible se consigue con la reunión de varios criterios, y si la identificación de los factores relevantes se hace a la luz de todo el marco teórico alrededor del constructo que se investiga (35,41,48,49).

Muestra para análisis de factores

Es importante tener presente que las validaciones de instrumentos se realizan para poblaciones con características específicas, por lo cual la validación es un proceso constante en el que se hacen modificaciones y adaptaciones a las escalas, de acuerdo con las particularidades culturales, sociales y lingüísticas de los participantes, y el contexto de aplicación, clínica o investigación (50-53). No obstante, es necesario contar con un tamaño de muestra adecuado y definir claramente las características de los participantes para contar con una aceptable representatividad que permita algún grado de extrapolación o generalización (29,42).

Por lo general, los estadísticos más liberales sugieren contar con al menos cinco personas por cada

elemento o reactivo (20,29,35). Muestras menores de 50 son inaceptables, y en el peor de los casos, se debe contar con al menos 100 participantes (20,29,35,36). Un grupo de 100 participantes es suficiente si se observan muchos reactivos con coeficientes altos en cada factor (43,49).

Cuando se pretende hacer análisis de factores con pocos ítems en cada factor o se esperan coeficientes muy bajos en cada uno de ellos, se sugiere incluir al menos diez personas por cada punto de la escala (20,35), como se recomienda para el cálculo de la consistencia interna cuando el instrumento consta de menos de diez ítems (3,18).

Sin embargo, los estadísticos más ortodoxos proponen el uso de muestras mucho más grandes, incrementar el número de participantes por ítems, con el ánimo de minimizar el error aleatorio que puede presentarse (27,41,54,55). Para ello, se diseñaron tablas que muestran el número de participantes por considerar para contar con una solución de factores con el menor error posible aceptable (54). Como norma general, estos autores indican la inclusión de al menos 20 personas por cada punto de la escala (27), como en otros modelos de regresión lineal múltiple (17). Algunos investigadores lo resumen de la siguiente forma: la muestra menor de 100 es insuficiente; alrededor de 200, modesta; de 300, buena; de 500, muy buena; y más de 1.000, excelente (55).

Conclusiones

El análisis de factores es un procedimiento útil durante el proceso de validación en salud mental para explorar o confirmar el número de factores (dimensionalidad) de una escala o instrumento de medición y, de esta forma, estimar parte de la validez. Existen los análisis de factores exploratorios y los confirmatorios, y diversos criterios para la extracción o retención final de los factores más importantes que subyacen en un número determinado de ítems. La combinación de criterios es la mejor opción de selección de factores. Se sugiere análisis de factores para grupo de elementos no superior a treinta, y en todos los casos se deben incluir por lo menos diez participantes por reactivo para tener una solución con el máximo error aceptable.

Referencias

1. Sánchez R, Echeverry J. Validación de escalas de medición en salud. *Rev Salud Pública*. 2004;6:302-18.
2. Prieto G, Delgado AR. Fiabilidad y validez. *Papeles Psicol*. 2010;31:67-74.
3. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess*. 2003;80:99-103.
4. Streiner DL. A checklist for evaluating the usefulness of rating scales. *Can J Psychiatry*. 1993;38:140-8.
5. Sánchez R, Gómez C. Conceptos básicos sobre validación de escalas. *Rev Colomb Psiquiatr*. 1998;27:121-30.
6. Kuder GF, Richardson MV. The theory of the estimation of test reliability. *Psychometrika*. 1937;2:151-60.
7. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull*. 1955;52:281-302.
8. Kaplan RM, Saccuzzo DP. *Pruebas psicológicas*. 6ta. edición. México: Thomson; 2006.
9. Cortina JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol*. 1993;78:98-104.
10. Raubenheimer J. An item selection procedure to maximise scale reliability and validity. *J Indust Psychol*. 2004;30:59-64.
11. Thompson B, Daniel LG. Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines (Editorial). *Educ Psychol Meas*. 1996;56:197-208.
12. Borsboom D, Mellenbergh GJ, Van Heerden J. The concept of validity. *Psychol Rev*. 2004;111:1061-71.
13. DiStefano C, Hess B. Using confirmatory factor analysis for construct validation: An empirical review. *J Psychoeduc Assess*. 2005;23:225-41.
14. Batista-Foguet JM, Coenders G, Alonso J. Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud. *Med Clin (Barc)* 2004;122(Supl. 1):21-7.
15. Kane MT. Current concepts in validity theory. *J Educ Meas*. 2001;38:319-42.
16. Lamprea JA, Gómez-Restrepo C. Validez en la evaluación escalas. *Rev Colomb Psiquiatr*. 2007;36:340-8.
17. Katz MH. *Multivariable analysis*. 2nd ed. Cambridge: Cambridge University Press; 2006.
18. Jaju A, Crask MR. The perfect design: optimization between reliability, validity, redundancy in scale items and response rates. *Am Market Assoc*. 1999;10:127-31.
19. Campo-Arias A, Oviedo HC. Propiedades psicométricas de una escala: la consistencia interna. *Rev Salud Pública*. 2008;10:831-9.
20. Streiner DL. Figuring out factors: the use and misuse of factor analysis. *Can J Psychiatry*. 1994;39:135-40.
21. Kahn JH. *Factor analysis in counseling psychology, research, training and practice: Principles, advances,*

- and application. *Counsel Psychol.* 2006;34:684-718.
22. Blacker D, Endicott J. Psychometric properties: concepts of reliability and validity. En: Rush AJ, Pincus HA, First MB, et al. *Handbook of psychiatric measures.* Washington: American Psychiatric Association; 2002 (CD-ROM).
 23. Roberts P, Priest H, Traynor M. Reliability and validity in research. *Nurs Stand.* 2006;20:41-5.
 24. Morgan GA, Gliner JA, Harmon RJ. Evaluating the validity of a research study. *J Am Acad Child Adolesc Psychiatry.* 1999;38:480-5.
 25. Rodríguez MA, Lopera J. Conceptos básicos en la validación de escalas en salud mental. *Revista CES Medicina.* 2002;16:31-9.
 26. Gorsuch RL. Exploratory factor analysis: its role in item analysis. *J Pers Assess.* 1997;68:532-60.
 27. Hogarty KY, Hines CV, Kromrey JD, et al. The quality of factor solutions in exploratory factor analysis: the influence of size sample, communalities, and over determination. *Educ Psychol Meas.* 2005;65:202-26.
 28. Costello AB, Osborne JW. Best practice in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pract Assess Res Eval.* 2005;10:7.
 29. Worthington RL, Whittaker TA. Scale development research. A content analysis and recommendations for best practices. *Counsel Psychol.* 2006;34:806-38.
 30. Bartlett MS. Test of significance in factor analysis. *Br J Psychol.* 1950;3:77-85.
 31. Kaiser HF. An index of factorial simplicity. *Psychometrika.* 1974;34:31-6.
 32. Reise SP, Waller NG, Comrey AL. Factor analysis and scale revision. *Psychol Assess.* 2000;12:287-97.
 33. STATA 11.0 for windows. College Station: StataCorp LP; 2011.
 34. SPSS for windows 19.0. Chicago: SPSS. Inc; 2011.
 35. Floyd FJ, Widaman KF. Factor analysis in the development and refinement of clinical assessment instruments. *Psychol Assess.* 1995;7:286-99.
 36. Henson RK, Roberts JK. Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educ Psychol Meas.* 2006;66:393-416.
 37. Russell DW. In search of underlying dimensions: The use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin.* *Pers Soc Psychol Bull.* 2002;28:1629-46.
 38. Jackson DL, Gillaspay JA, Purc-Stephenson R. Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychol Methods.* 2009;14:6-23.
 39. Zwick WR, Velicer WF. Factors influencing four rules for determining the number of components to retain. *Multiv Behav Res.* 1988;17:253-69.
 40. Ferré L. Selection of components in principal component analysis: A comparison of methods. *Comput Stat Data Anal.* 1995;19:669-82.
 41. Coste J, Bouée S, Ecosse E, et al. Methodological issues in determining the dimensionality of composite health measures using principal component analysis: Case illustration and suggestions for practice. *Qual Life Res.* 2005;14:641-54.
 42. Fabrigar LR, Weneger DT, MacCallum RC, et al. Evaluating the use of exploratory analysis in psychological research. *Psychol Methods.* 1999;4:272-99.
 43. Norman GR, Streiner DL. *Bioestadística.* Madrid: Mosby/Doyma; 1996.
 44. Stevens J. Power of the multivariate analysis of variance tests. *Psychol Bull.* 1986;88:728-37.
 45. Polit D, Hungler B. *Investigación científica en ciencias de la salud.* 5ª edición. México: Interamericana-McGraw-Hill; 2000.
 46. Macía F. *Validez de los tests y análisis factorial: Nociones generales.* Cienc Trab. 2010;12:276-80.
 47. Cattell RB. The scree test for the number of factors. *Multiv Behav Res.* 1966;1:245-76.
 48. Hurley AE, Scandura TA, Schriesheim CA, et al. Exploratory and confirmatory factor analysis: guidelines, issues, and alternatives. *J Organ Behav.* 1997;18:667-83.

49. Hayton JC, Allen DG, Scapelo V. Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organ Res Methods*. 2004;7:191-205.
50. Strauss E, Spreen O, Hunter O. Implications of test revisions for research. *Psychol Assess*. 2000;12:237-44.
51. Silverstein ML, Nelson LD. Clinical and research implications of revising psychological tests. *Psychol Assess*. 2000;12:298-303.
52. Rubio-Stipec M, Hicks MHR, Tsuang MT. Cultural factors influencing the selection, use, and interpretation of psychiatric measures. En: Rush AJ, Pincus HA, First MB, et al. (editors). *Handbook of psychiatric measures*. Washington: American Psychiatric Association; 2002 (CD-ROM).
53. Bland JM, Altman DG. Validating scales and indexes. *BMJ*. 2002;324:606-7.
54. Charter RA. Sample size requirement for precise estimates of reliability, generalizability, and validity coefficients. *J Clin Exp Neuropsychol*. 1999;21:559-66.
55. MacCallum RC, Widaman KF, Zhang S, et al. Sample size in factor analysis. *Psychol Methods*. 1999;4:84-99.

Conflictos de interés: Los autores manifiestan que no tienen conflictos de interés en este artículo.

Recibido para evaluación: 2 de diciembre de 2011

Aceptado para publicación: 3 de abril de 2012

Correspondencia

Adalberto Campo-Arias

Instituto de Investigación del Comportamiento Humano

Calle 58 No. 5-24, oficina 202

Bogotá, Colombia

campoarias@comportamientohumano.org