

Investigaciones y productos CID



UNIVERSIDAD NACIONAL DE COLOMBIA
SEDE BOGOTÁ
FACULTAD DE CIENCIAS ECONÓMICAS

N° 18

ISSN 2027-8780

Mayo 2014

Prueba escrita de conocimientos para evaluar aspirantes al cargo de juez de la república: propiedades psicométricas y estructura factorial en una muestra colombiana

Reliability and Factorial Structure of a Knowledge Test to Assess Candidates for the Position of Judge of the Republic of Colombia

Fabian González

CID Centro de
Investigaciones
para el Desarrollo

PRUEBA ESCRITA DE CONOCIMIENTOS PARA EVALUAR ASPIRANTES AL CARGO DE JUEZ DE LA REPÚBLICA: PROPIEDADES PSICOMÉTRICAS Y ESTRUCTURA FACTORIAL EN UNA MUESTRA COLOMBIANA

Fabian González¹

Resumen: Este artículo presenta los resultados del estudio sobre las propiedades psicométricas y estructura factorial de una prueba de conocimientos aplicada a los aspirantes al cargo de Juez de la República en una muestra colombiana. Se examina la fiabilidad y la validez de constructo, y se analiza el modelo mediante análisis factorial confirmatorio. Los resultados indican una elevada fiabilidad, tanto para las dimensiones como la evaluación global del desempeño. Un análisis de componentes principales indica la presencia de factores en los que cargan las valoraciones de los superiores inmediatos y los criterios objetivos externos respectivamente. El análisis factorial confirmatorio indica que el desempeño puede representarse mediante un modelo jerárquico únicamente para la prueba de conocimientos específicos. Las implicaciones de los resultados se comentan en la discusión.

Palabras clave: pruebas escritas, psicometría, validez de constructo, análisis factorial.

Código JEL: M54

¹ Psicólogo de la Universidad Nacional de Colombia contacto:
flgonzalezr@unal.edu.co

RELIABILITY AND FACTORIAL STRUCTURE OF A KNOWLEDGE TEST TO ASSESS CANDIDATES FOR THE POSITION OF JUDGE OF THE REPUBLIC OF COLOMBIA

Abstract: This article presents the results of the study on the factorial structure and psychometric properties of a knowledge test applied to the Colombian applicants for the post of Judge. The study examines the reliability and construct validity, and analyzes the model using confirmatory factor analysis.

Keywords: selection tests, psychometrics, construct validity, factor analysis.

JEL Classification : M54

En esta colección se publican ensayos; resultados de investigación previos o concluidos; documentos producto de consultorías; reflexiones de investigadores adscritos al Centro de Investigaciones para el Desarrollo-CID; documentos de investigadores invitados a eventos realizados por el CID; y, en general, documentos caracterizados por su rigurosidad científica y pertinentes a un contexto específico o coyuntural.

Investigaciones y productos – CID **ISSN 2027-8780**

Investigaciones y productos FCE-CID puede ser consultada en el portal virtual:

<http://www.fce.unal.edu.co/publicaciones/>

Director Centro Editorial-FCE
Álvaro Zerda Sarmiento

Equipo Centro Editorial-FCE
Nadeyda Suárez Morales
Maria del Pilar Ducuara López
Jesús David Suárez Sosa

Contacto: Centro Editorial FCE-CID
Correo electrónico: publicac_fcebog@unal.edu.co

Este documento puede ser reproducido citando la fuente. El contenido y la forma del presente material es responsabilidad exclusiva de sus autores y no compromete de ninguna manera al Centro de Investigaciones para el Desarrollo CID, ni a la Facultad de Ciencias Económicas, ni a la Universidad Nacional de Colombia.

Rector

Ignacio Mantilla Prada

Vicerrector General

Jorge Iván Bula Escobar

Facultad de Ciencias Económicas

Decano

José Guillermo García Isaza

Vicedecano

Rafael Suárez

Centro de Investigaciones para El Desarrollo CID

Director

Jorge Armando Rodríguez

Subdirectora

Vilma Narváez

Contenido

Resumen.....	2
Abstract.....	3
Introducción.....	6
Método.....	10
Participantes.....	10
Instrumentos.....	10
Procedimientos.....	10
Resultados.....	10
Análisis factorial confirmatorio.....	11
Análisis de fiabilidad.....	13
Discusión.....	13
Referencias.....	15

Introducción

Quienes tienen la responsabilidad de vincular personal en sus organizaciones deben tener información adecuada y suficiente para tomar las decisiones correctas. La apropiada selección de personal dependerá de la definición de los propósitos de la evaluación, la precisión en la definición del perfil del cargo, la utilización de instrumentos confiables para evaluar las diferencias individuales y la validez de los procedimientos de evaluación utilizados.

El principal objetivo de los instrumentos de medición y particularmente de las pruebas escritas de selección de personal es el de obtener información acerca de las competencias, aptitudes, inteligencia y rasgos de personalidad del aspirante, así como predecir su probable éxito en el desempeño del cargo.

Una prueba debe ser una medida objetiva y estandarizada de la conducta, las habilidades, la inteligencia o los rasgos de personalidad del individuo. En primer lugar, la muestra de conducta debe ser suficientemente grande y representativa de lo que pretende medir para poder generalizar y predecir a partir de sus resultados. En segundo lugar, al exigir que la prueba sea estándar, se requiere que la serie de preguntas o tareas deben administrarse uniformemente. Por último, todas las pruebas psicométricas deben tener las características de confiabilidad, validez y estandarización. (Anastasi, 1974 y Hernández y cols, 2006). La validez hace referencia a la capacidad del instrumento de medir el constructo que pretende cuantificar y la confiabilidad, a la propiedad de mostrar resultados similares en repetidas mediciones (Campo, 2008).

Algunas propiedades psicométricas se pueden calcular con la aplicación del instrumento en una sola ocasión y en una muestra o población objeto de estudio (Hebson, 2001). Las características de un instrumento que se pueden explorar con una única aplicación son la consistencia interna (homogeneidad), la dimensionalidad (estructura factorial o interna) asociada a la validez del constructo (Blacker, 2002).

El presente artículo se centra en la estimación de la consistencia interna y la dimensionalidad de la prueba escrita de conocimientos aplicada a los aspirantes al cargo de Juez de la República en

una muestra colombiana.

La confiabilidad relacionada con la consistencia interna se refiere al grado en que las preguntas o ítemes que conforman la prueba escrita se correlacionan entre ellos, es decir la magnitud en que miden el mismo constructo. Si los puntos que componen una escala teóricamente miden el mismo constructo deben mostrar una alta correlación, es decir, la escala debe mostrar un alto grado de homogeneidad (Hebson, 2001 y Cortina, 1993).

Aunque habitualmente se acepta que la consistencia interna de una escala puede estar en el rango entre 0 y 1, teóricamente, los valores podrían oscilar entre -1 y 1 (Krus, 1993). Valores de 1 indicarían una correlación positiva entre ítemes cercanas a la perfección; valores de 0 , ninguna correlación; y coeficiente de -1 , una correlación negativa entre los ítemes.

Por otra parte, la validez de constructo incluye la relevancia y representatividad de los contenidos, así como las relaciones con los criterios, ya que ambas dan significado a las puntuaciones de las pruebas (Martínez Arias, 1995, p.335). Por consiguiente, la validez de constructo puede ser conceptualizada en términos de un proceso científico de contrastación de hipótesis, donde participa tanto lo empírico como los juicios racionales: las hipótesis serían las inferencias realizadas a partir de las puntuaciones obtenidas en las pruebas y la validación el proceso de acumulación de evidencias que apoyen dichas inferencias, usos o interpretaciones de las puntuaciones de la prueba (Messick, 1989).

Entre los procedimientos o técnicas estadísticas utilizados para la contrastación de la validez de constructo predomina el Análisis Factorial. En general, se puede decir que es la técnica más utilizada para estimar la validación de constructo.

Cuando se analiza la estructura estadística de las respuestas a una prueba escrita, generalmente se utiliza la técnica del análisis factorial exploratorio. En este tipo de técnica no se requiere establecer *a priori* cuál es la estructura de los datos; por el contrario, son los propios datos, en función de unos criterios empíricos los que evidencian su estructura. Este tipo de técnica es de gran utilidad cuando se desconoce de antemano qué tipo de estructura puede esperarse de las respuestas de los sujetos evaluados.

Una situación, sin embargo, que no es deseable ya que el investigador ha debido definir

operativamente un constructo en función de una teoría sólida y, por tanto, es muy probable que pueda anticipar la estructura del constructo. La mayoría de los investigadores, en consecuencia, utilizan esta técnica para confirmar empíricamente la estructura conceptual que han establecido de antemano y para conocer el papel que juega cada elemento en el conjunto global de esa estructura (Herrero, 2010).

Existe otro acercamiento que ofrece mayor flexibilidad al establecer sus hipótesis sobre la estructura del constructo (Bentler, 2007). Esta técnica, denominada análisis factorial confirmatorio, permite contrastar un modelo construido con antelación, en el que se establece *a priori* el conjunto total de las relaciones entre los elementos que lo configuran. A diferencia del análisis factorial exploratorio, en el análisis factorial confirmatorio se supone que el investigador es capaz de estimar *a priori* la estructura de los datos -preferiblemente en función de una teoría bien establecida- y sólo precisa confirmar que esa estructura puede también obtenerse empíricamente. Técnicamente, las dos estrategias persiguen un único objetivo: “explicar las covarianzas o correlaciones entre un conjunto de variables observadas o medidas a través de un conjunto reducido de variables latentes o factores” (Bollen, 1989).

Desde el punto de vista del Análisis Factorial Confirmatorio, la puntuación de cada sujeto en cada ítem está generada por una variable no observada (el Factor Latente) que explica la variabilidad de las puntuaciones en el ítem. Previsiblemente, el Factor Latente nunca explicará de forma totalmente satisfactoria la variabilidad de las respuestas del ítem. A esta parte no explicada por el factor se le denomina error de medida (Schreiber y cols., 2006).

Una ventaja inicial del análisis confirmatorio frente al exploratorio es que separa de la varianza de cada ítem, la parte de la varianza explicada por el factor y la parte que no explica el factor para, posteriormente, diferenciar ambas variables y calcular sus coeficientes y varianzas por separado. De este modo, una vez identificado el error, se trabaja sólo con la parte de la puntuación del ítem que se considera representa a los conocimientos evaluados.

Por otra parte, el espacio teórico que posibilita el análisis confirmatorio en la interpretación del error es muy flexible. Así, la parte de varianza que no explican determinados factores puede ocasionarse por diferentes circunstancias. Entre las circunstancias que explican esta variación se pueden distinguir aquellas que tienen que ver con la naturaleza del cuestionario y aquellas que tienen que ver con la naturaleza del concepto que intenta representar ese cuestionario (Cole y cols., 1993).

Las circunstancias que tienen que ver con la naturaleza del cuestionario tienden a explicar el error en función de las características del cuestionario o del ítem. Por ejemplo, Bollen (1989) ha señalado que el error puede deberse a que los indicadores proceden de la misma fuente o a que existen sesgos en las respuestas al cuestionario. Por otra parte, en las circunstancias que tienen que ver con la naturaleza del concepto se ubica la existencia de un espacio para la teoría en la interpretación de los términos de error. El término de error puede representar aquella parte del ítem que no tiene que ver con el constructo medido pero que contiene información valiosa para el

investigador.

La lógica que subyace en el análisis confirmatorio es la de eliminar esa fuente de variación en la configuración del factor en el que satura el ítem pero, si es preciso, utilizarla para explicar algunas relaciones en el modelo (Bentler, 1993). En cualquier caso, el análisis factorial confirmatorio siempre requiere de la existencia de un marco teórico articulado que sirva de base para la elaboración de un modelo cuya contrastación empírica se está analizando. Por esta razón, es necesario que el modelo muestre un ajuste adecuado a los datos y, además, que los parámetros que componen el modelo muestren la dirección y significación previstas.

En términos generales, los dos tipos de procedimientos corresponden con las dos grandes aproximaciones que se suelen llevar a cabo para la definición de los constructos: la aproximación inductiva o exploratoria y la deductiva o confirmatoria. La aproximación inductiva o exploratoria supone delimitar un número amplio de indicadores que supuestamente miden el constructo, seguidamente estos indicadores son analizados mediante el análisis factorial exploratorio para buscar patrones de relación entre los indicadores y, finalmente, *a posteriori*, denominar esos patrones de relación para definir el constructo. En la aproximación deductiva o confirmatoria, por otra parte, se comienza con un modelo teóricamente aceptado y asumido para describir y/o explicar los datos empíricos. La construcción del modelo está basada en una información *a priori* sobre la naturaleza de la estructura de los datos, o bien en una teoría sustantiva en el campo de trabajo del que se trate. El procedimiento permite una valoración de la correspondencia de semejanza entre las características del concepto planteado y los datos obtenidos sobre dicho concepto a través de sus indicadores, es decir, la evaluación de la correspondencia teórica entre el constructo y el dato (Schmitt, 1995).

Es de suponer que esta manera de proceder es correcta en sí misma y se ajusta a los criterios de validez de constructo ya señalados, es decir, asumida *a priori* una estructura teórica, se obtiene una muestra representativa de la población sobre la cual se aplica el análisis factorial. Si la estructura obtenida es coincidente con la estructura teórica se ha confirmado nuestro modelo teórico.

El consejo Superior de la Judicatura ha decidido realizar pruebas objetivas dentro de los procesos de ingreso con el fin de garantizar igualdad de oportunidades para todos los aspirantes, asegurando transparencia y blindando el proceso de las influencias que durante muchos años incidieron en el nombramiento de empleados y funcionarios. En este modelo meritocrático se valoran los saberes o dominios adquiridos por los aspirantes que deciden ingresar a la carrera judicial, estimula la participación masiva y fortalece las políticas de gestión con un enfoque de calidad y transparencia al interior del Consejo Superior de la Judicatura. Las pruebas, diseñadas y construidas por docentes de la Universidad Nacional, fortalecen la administración de justicia al permitir el ingreso y la continuidad de funcionarios con altas calidades profesionales. Estos

instrumentos garantizan un proceso basado en información confiable y verificable.

Método

Participantes

La muestra definitiva la componen 4.712 aspirantes al cargo de Juez Civil con Competencias Laborales, evaluados simultáneamente en 28 ciudades a través de un instrumento estructurado.

Instrumento

El instrumento de evaluación es un cuestionario de 150 preguntas cerradas, diseñadas y elaboradas por docentes de la Facultad de Derecho de la Universidad Nacional de Colombia, quienes definieron los temas específicos y el nivel de dificultad esperado. Todas las preguntas se sometieron a un proceso de validación por jueces expertos que conocen del quehacer de los Jueces Civiles del Circuito y por tanto, esta validez está dada por la metodología de construcción de las pruebas. El instrumento tenía dos componentes: uno general con 50 preguntas y otro específico con 100 preguntas. El componente general evaluó conocimientos en Ley Estatutaria de Administración de Justicia y Sus Reformas, Interpretación Jurídica, Derecho Constitucional, Parte General del Derecho Procesal y Derecho Probatorio. El componente específico evaluó conocimientos en Derecho Civil, Derecho Comercial, Derecho de Familia, Derecho Agrario y Derecho Laboral.

Procedimiento

Una vez obtenidas las respuestas al cuestionario aplicado a nivel nacional se realizaron los análisis para establecer las características psicométricas, teniendo como metas la estimación de la fiabilidad, la estructura interna y la validación de constructo del modelo mediante un análisis factorial confirmatorio. Se aplicó la técnica de análisis factorial con Máxima Verosimilitud (ML) como procedimiento de extracción de los factores y el procedimiento VARIMAX u OBLIMIN, los cuales definen la rotación según el correspondiente modelo analizado. Para ello se utilizó el módulo PROC FACTOR del programa SPSS v20 (Spss INC, 1994). En las páginas que siguen se describe el desarrollo y los resultados de dicho estudio.

Resultados

Al eliminar las respuestas omitidas, la muestra efectiva considerada en estos análisis fue de 4.712 aspirantes. Se discriminaron los ítems y quedaron aquellos que obtuvieron puntuaciones en la correlación ítem-total mayores a 0,30 (Kline, 1997; Tabachnick y Fidell, 1996) para los análisis posteriores que se detallan a continuación:

Análisis factorial confirmatorio

Aunque la técnica estadística lleva el nombre de confirmatorio, la estrategia de análisis es exploratoria. En primer lugar se analizó si la matriz de correlaciones para los ítems de cada escala era una matriz apropiada para realizar un análisis factorial. La prueba de esfericidad de Bartlett (1345.68; $p < .001$ para la prueba general y 8892.47; $p < .001$ para la prueba específica) indicó que los ítems eran independientes únicamente para la prueba específica. Los coeficientes de Kaiser-Meyer-Olkin ofrecieron valores de .83 y .76 respectivamente, indicando que las correlaciones entre parejas de ítems pueden ser explicadas por los restantes ítems seleccionados.

Se llevó a cabo la validez factorial del instrumento con diferentes métodos de extracción y de rotación con el objeto de identificar la solución más adecuada con relación al contenido de los ítems. Con el propósito de identificar la estructura conceptual más clara para los cinco factores sugeridos por los constructores de la prueba general y los dos factores de la prueba específica, se eligió una rotación oblimín, debido al elevado nivel de correlación entre las dimensiones.

Se utilizó como método final el de Máxima Verosimilitud (ML) ajustando la solución a los factores que obtuvieron valores propios (eigenvalues) superiores a 1 (Tabachnick et al., 1996).

Conjuntamente los factores explicaron el 12,7% de la varianza total del modelo para la prueba general y el 42.6% para la prueba específica. Esto se explicaría porque a diferencia de la prueba general, los contenidos evaluados en la prueba específica están compuestos por temas delimitados teórica y conceptualmente para los cuales los aspirantes se han preparado y han obtenido la mayor parte de su experiencia laboral.

Los 17 factores o dimensiones encontradas se relacionan con los subtemas definidos para las dos grandes áreas de conocimientos evaluadas: Derecho Civil y Derecho Laboral. Por motivos de confidencialidad en la información, este estudio no puede revelar el contenido de los ítems que conformarían dichos factores.

Tabla 1. Análisis factorial para explorar los posibles dominios de los ítems que conforman la prueba de conocimientos específicos. Varianza total explicada

Factor	Autovalores iniciales		
1	2,936	5,871	5,871
2	1,560	3,120	8,992
3	1,373	2,746	11,737
4	1,320	2,640	14,378
5	1,236	2,472	16,850
6	1,169	2,339	19,189
7	1,156	2,312	21,501
8	1,143	2,286	23,787
9	1,105	2,209	25,996
10	1,095	2,191	28,187
11	1,064	2,129	30,315
12	1,054	2,107	32,422
13	1,047	2,093	34,516
14	1,040	2,080	36,596
15	1,021	2,041	38,637
16	1,014	2,028	40,665
17	1,005	2,011	42,675
18	,999	1,998	44,673
19	,991	1,982	46,655
20	,979	1,957	48,612
21	,972	1,944	50,557
22	,969	1,937	52,494
23	,962	1,925	54,418
24	,955	1,910	56,328
25	,937	1,873	58,201
26	,928	1,856	60,057
27	,926	1,852	61,909
28	,919	1,837	63,746
29	,908	1,815	65,561
30	,901	1,803	67,364
31	,897	1,795	69,159
32	,892	1,784	70,943
33	,880	1,761	72,703
34	,871	1,741	74,445
35	,865	1,730	76,175
36	,859	1,719	77,894
37	,857	1,713	79,607
38	,847	1,694	81,300
39	,840	1,680	82,981
40	,829	1,658	84,639
41	,818	1,635	86,274
42	,809	1,619	87,893
43	,795	1,589	89,483
44	,783	1,566	91,049
45	,777	1,553	92,602
46	,762	1,524	94,127
47	,757	1,514	95,640
48	,744	1,487	97,128
49	,726	1,451	98,579
50	,711	1,421	100,000

Método de extracción: Máxima verosimilitud.

Análisis de fiabilidad

El índice de fiabilidad evaluado a través del alpha de Cronbach, indica que la prueba en su totalidad tiene una muy buena consistencia interna ($\alpha = .88$). En cuanto a los ítems de la escala, no hay ningún ítem cuya eliminación aumente la consistencia interna de la escala. Sin embargo, la prueba general tiene un índice relativamente menor y sugiere una revisión de sus componentes en relación con la pertinencia y la población objetivo.

Tabla x. Índice de fiabilidad de las escalas de la prueba escrita

	P. General	P. Civil	P. Laboral
number of subtest items:	50	60	40
minimum possible score:	0	0	0
maximum possible score:	50	60	40
index of reliability:	0,78	0,98	0,96
standard error of measurement:	2,64	2,93	2,44

En relación con la calidad de los reactivos, específicamente con su capacidad de discriminación, se observan valores que varían entre .20 y .40. Estos son índices adecuados teniendo en cuenta la naturaleza de las preguntas y la población objeto de evaluación.

Discusión

En las últimas décadas se han construido instrumentos para medir o evaluar el conocimiento general y específico en procesos de selección de aspirantes a cargos en el sector público, sin que en la mayoría de los casos se haya estudiado las propiedades psicométricas de la herramienta luego de su aplicación. Disponer de un instrumento válido y confiable para la población objetivo es de gran importancia para obtener información que permita predecir exitosamente el desempeño dentro de los puestos de trabajo convocados. AL respecto, es importante aclarar que

la muestra de participantes con la que se realizó la valoración del cuestionario no se obtuvo mediante una técnica probabilística, ya que el objetivo fundamental era efectuar la estimación de las propiedades psicométricas del instrumento. En consecuencia, los resultados obtenidos no pueden generalizarse o asegurarse que tienen validez externa.

Los hallazgos del presente estudio indican que el instrumento de conocimientos específicos aplicado a nivel nacional y con propósitos de selección de funcionarios para la Rama Judicial se adaptó consistentemente entre la población de aspirantes evaluados. Se logró obtener, a través de un análisis factorial, subescalas más específicas conformadas por ítems consistentes conceptual y estadísticamente. Los resultados claramente mostraron una discriminación en los puntajes obtenidos en el componente civil y laboral a diferencia de los obtenidos en el componente general. Esta distinción hace que se identifiquen estadísticamente las dimensiones elaboradas para la prueba específica.

La prueba general debe someterse a un proceso de revisión de pertinencia y relevancia. Los resultados obtenidos no son satisfactorios en relación con la delimitación de los factores inicialmente diseñados y por lo contrario, el comportamiento refleja que un modelo monofactorial es el más adecuado para representar los datos.

El comportamiento de los datos refleja que en la prueba general se identifica un factor principal referido a conocimientos globales sobre el quehacer del juez civil, sin que se permita identificar con claridad las dimensiones que lo conforman. Es al parecer una dimensión muy general y ambigua a la vez.

Para la prueba específica se encontraron evidencias de un buen ajuste a través de áreas más delimitadas referidas a los temas que componen las preguntas para Derecho Civil, Comercial, Agrario y Laboral.

Si bien los resultados de este estudio arrojaron 17 dimensiones alrededor de los conocimientos específicos y cuya consistencia interna fue elevada, no se encontraron evidencias de las dos subescalas claramente delimitadas y originalmente propuestas por los constructores de la prueba. Una posible explicación se debe a que los temas en derecho civil pueden ser de un ámbito aplicable a otras áreas del derecho evaluadas en el mismo instrumento. Otra explicación tiene que ver con las variaciones en los resultados obtenidos por los especialistas en Derecho Civil y por los Especialistas en Derecho Laboral. Estas dos áreas son de distinto contexto laboral en la práctica, pero al ser evaluadas en un mismo cuestionario, puede generar un sesgo en relación con la formación y experiencia laboral del aspirante.

En resumen, la importancia de este estudio radica en corroborar las dimensiones identificadas en los cuestionarios aplicados en procesos de selección masivos para proveer cargos en el sector

público en Colombia, lo que brindará mayores evidencias de la estructura teórica del instrumento y su aplicabilidad en una población determinada. Un siguiente objetivo consiste en correlacionar los factores evaluados con los índices de desempeño esperados en los puestos de trabajo, ya que arrojará mayor conocimiento sobre el carácter predictivo de los instrumentos de selección contruidos por expertos en las diferentes áreas del conocimiento.

Finalmente, los resultados encontrados sugieren la realización de un análisis factorial con un muestreo probabilístico y el posterior proceso estandarización y obtención de baremos de la escala al contexto laboral de la rama judicial a nivel nacional.

Referencias

Anastasi, A. (1974). *Psychological testing* (2a. edición). New York: Macmillan Hernández R, Fernández-Collado C, Baptista P. (2006). *Metodología de la investigación*. Cuarta edición. México: McGraw-Hill Interamericana.

Campo, A, Oviedo, H. (2008). Propiedades Psicométricas de una Escala: la Consistencia Interna. *Rev. salud pública*. 10 (5):831-839

Hebson RK. (2001). Understanding internal consistency reliability estimates: a conceptual primer on coefficient alpha. *Meas Eval couns Dev*; 34: 177-189.

Blackler D, Endicott J (2002). *Psychometric properties: concepts of reliability and validity*; 2002.

Cortina JM. (1993). What is coefficient alpha? An examination of theory and applications. *J Appl Psychol*; 78: 98-104.

Kline, P. (1997). *An easy guide to factor analysis*. Routledge: Londres.

Krus DJ, Helmstadter GC. (1993). The problem of negative reliabilities. *Educ Psychol Meas*; 53:

643-650.

Martínez-Arias, M.R. (1995). *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid: Síntesis.

Messick, S. (1989). Validity. The specific ation and development of tests of achievement and ability. En R. L. Lino (Ed.), *Educational Measurement (3th edition)*. Washington, DC: American Council on Education.

Herrero, J. (2010). El Análisis Factorial Confirmatorio en el estudio de la Estructura y Estabilidad de los Instrumentos de Evaluación: Un ejemplo con el Cuestionario de Autoestima CA-14. *Intervención Psicosocial* Vol. 19, n. 3 - pp. 289-300

Bentler, P. M. (2007). Can scientifically useful hypotheses be tested with correlations? *American Psychologist* , 62, 772-782.

Bollen, K. A. (1989). *Structural equations with latent variables* . New York: John Wiley y Sons.

Cole, D. A., Maxwell, S. E., Arvey, R. y Salas, E. (1993). Multivariate group comparisons of variable systems: MANOVA and structural equation modeling.

Psychological Bulletin, 114, 174-184.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A. y King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research*, 99, 323-337.

Bentler, P. M. (1993). *EQS. Structural Equations Program Manual*. Los Angeles: BMDP Statititcal Software Inc.

Schmitt, F.F. (1995). Truth: A primer. Boulder, Colorado:Westviw Press.

Sppss Inc., M. J. Norusis (1994). SPSS/PC+ for IBM=PC/XT/AT. Chicago, IL: SPSS Inc.

Tabachnick, G.T. y Fidell, L.S. (1996). Principal components and factor analysis: Using multivariate statistics (3ª edición). Nueva York: Harper Collins College.

