

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

Confirmatory Factor Analysis in the study of the Structure and Stability of Assessment Instruments: An example with the Self-Esteem Questionnaire (CA-14)

Juan Herrero

Universidad de Oviedo - España

Resumen. Este trabajo presenta un estudio de la estructura factorial y la estabilidad del Cuestionario de Autoestima CA-14 utilizando la técnica del Análisis Factorial Confirmatorio. El trabajo pretende ilustrar y guiar en las posibilidades que ofrece esta técnica, prestando especial atención a los requisitos que deben cumplir los datos, los métodos de estimación sugeridos en la literatura científica, los índices de ajuste más adecuados para evaluar los modelos y otras circunstancias que se deben tener en cuenta a la hora de estimar modelos de Análisis Factorial Confirmatorio. En el trabajo se presentan además diversas estrategias metodológicas en la implementación de esta técnica: correlación de errores residuales, imposición de restricciones o equivalencias en los parámetros de un modelos, modelos multigrupo, etc.

Palabras clave: análisis factorial confirmatorio, análisis multigrupo, autoestima, población general adulta.

Abstract. This research presents a study of the factor structure and temporal stability of the Self-Esteem Questionnaire (CA-14) using Confirmatory Factor Analysis (CFA) techniques. The paper pretends to offer a potential guide to researchers, paying special attention to data requirements, estimation methods suggested in the literature, recommended fit indices and other circumstances that need to be taken into account when estimating CFA models. Also, various methodological strategies are shown during the implementation of CFA models: correlated errors, use of parameter constraints, multigroup analysis, etc.

Keywords: Confirmatory Factor Analysis, multigroup analysis, self-esteem, adult general population.

Introducción

Cuando se analiza la estructura estadística de las respuestas a un cuestionario, es común utilizar la técnica del análisis factorial exploratorio. En este tipo de técnica el investigador no precisa establecer *a priori* cuál es la estructura de los datos; son los propios datos, en función de unos criterios empíricos, quienes muestran su estructura. Este tipo de técnica es muy útil cuando el investigador desconoce de antemano qué tipo de estructura puede esperar de las respuestas de los sujetos que componen la muestra. Una situación, sin embargo, que no es nada deseable ya que el investigador ha debido operativizar adecuadamente 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, por tanto, 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 (varianza total explicada por los factores, varianza explicada de cada factor y saturación de los ítems en los factores, fundamentalmente).

Análisis factorial exploratorio vs confirmatorio

Existe otro acercamiento que permite al investigador mayor flexibilidad al establecer sus hipótesis sobre la estructura del constructo (Bentler, 2007). Esta técnica, el análisis factorial confirmatorio, permite contrastar un modelo construido con antelación, en el que el investigador establece *a priori* el conjunto total de las relaciones entre los elementos que lo configuran. A diferencia del factorial exploratorio, en el factorial confirmatorio se supone que el investigador es capaz de aventurar *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, ambas estrategias persiguen un único objetivo: “explicar las covarianzas o correlaciones entre un conjunto de variables observadas

La correspondencia sobre este artículo puede dirigirse a Juan Herrero. Despacho 211. Departamento de Psicología. Plaza Feijoo s/n. 33003 Oviedo. E-mail: olaizola@uniovi.es

o medidas a través de un conjunto reducido de variables latentes o factores” (Bollen, 1989, p. 226).

Para ilustrar la lógica subyacente al Análisis Factorial Confirmatorio es preciso introducir previamente el concepto de Factor Latente que se utiliza en el Análisis Confirmatorio. Utilizaré un Cuestionario de Autoestima de 14 ítems, que es una versión abreviada de un cuestionario publicado previamente (CA-17, Gracia, Herrero y Musitu, 2002) para ejemplificar los conceptos fundamentales de esta técnica (ver también Herrero y Gracia, 2007, para un ejemplo de Análisis Confirmatorio con diferentes variables). Esta versión abreviada evalúa la autoestima en población adulta en torno a cuatro aspectos: físico, emocional, familiar y social.

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 (E).

Supongamos, en el caso más simple, que la estructura del CA-14 refleja un único factor de autoestima general (SELF en la figura 1). Este único factor da cuenta de las relaciones entre los 14 ítems; es decir, de sus correlaciones o covarianzas. Por tanto, la parte que estos 14 ítems comparten entre sí puede atribuirse a la existencia de una estructura latente -en este caso, configurada en torno a un sólo factor- que explica en

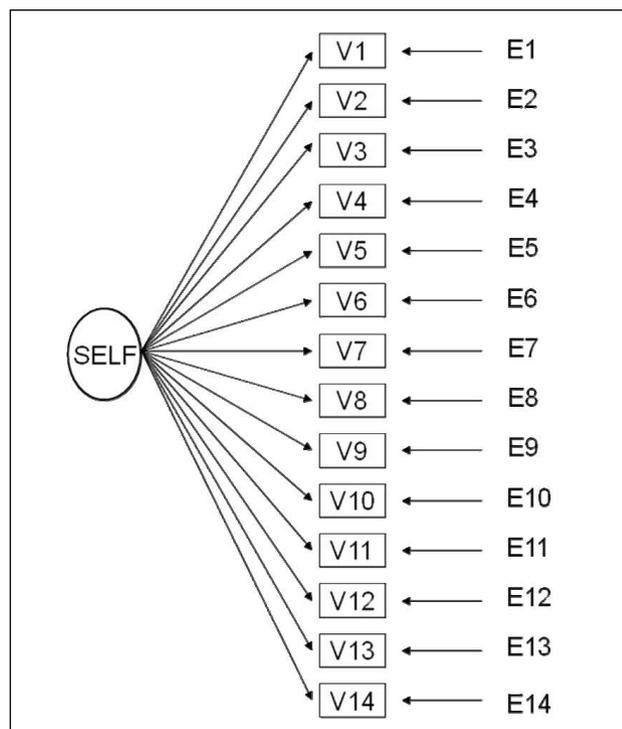
buena parte (varianza explicada) las correlaciones entre los ítems. En el análisis exploratorio, los factores dependen de la comunalidad entre los ítems, que son básicamente la cantidad de varianza explicada que una combinación de los restantes ítems ejerce sobre un ítem en particular; en este caso, los aspectos comunes que se observan en los 14 ítems se representan gráficamente por una flecha que va desde SELF a cada una de las variables que componen este factor. En este caso, todos los ítems comparten algo de la autoestima global y por tanto todos tienen una flecha que los conecta con el factor. La dirección de la flecha indica que es el factor quien genera las puntuaciones en el ítem y no viceversa. Ello quiere decir que cada ítem mide adecuadamente el factor al que pertenece en el grado en que su relación con el factor es mayor. Así, la varianza total de las respuestas a un ítem -digamos, el ítem 1- está compuesta por la varianza que comparte con el factor SELF -la flecha entre ítem V1 y SELF- y la varianza que no comparte con el factor -la flecha entre E1 y el ítem V1. La versión estandarizada de uno y otro valor se conoce como saturación y error, respectivamente. La suma de todas las saturaciones al cuadrado -la correlación al cuadrado de cada ítem con el factor- indica la cantidad total de varianza explicada de cada factor (el eigenvalue de cada factor). Alternativamente, uno menos el error al cuadrado indica la proporción de varianza no explicada por el factor.

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, trabaja sólo con la parte de la puntuación del ítem que se considera representa a la autoestima. Este tipo de factores están por tanto libres de error de medida.

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 explica el factor general SELF del ítem V1 - E1-, puede deberse a una pluralidad de 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.

1. Circunstancias que tienen que ver con la naturaleza del cuestionario. Este tipo de circunstancias representa la conceptualización más restrictiva del error y tiende 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. Así, mostrar acuerdo-desacuerdo en ítems formulados de forma negativa

Figura 1, Modelo 1: Un solo factor de autoestima para los 14 ítems 1



puede conllevar cierto error en la comprensión del ítem. El hipotético ítem, *no me disgusta que hablen mal de mí*, no está exento de dificultad de interpretación y, por tanto, proporciona un espacio para el error; esto es, aquella parte de la respuesta al ítem que no contesta realmente al ítem. Este error correlacionará, probablemente, con el error en las respuestas a otro tipo de ítems formulados de forma similar; por ejemplo, mis relaciones sociales no son insatisfactorias (acuerdo/desacuerdo). De esta forma, la estimación de la correlación de los errores puede incrementar la capacidad del modelo para reflejar los datos reales, identificando además de forma más precisa las fuentes de variación ajenas a los factores. En otras palabras, controla la variación en las puntuaciones producida por el error de medida.

2. Circunstancias que tienen que ver con la naturaleza del concepto. En general, siempre existe 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. Por ejemplo, un hipotético ítem de consumo de sustancias como *fumo marihuana con mis amigos*, contiene información tanto sobre el consumo de sustancias como sobre la actividad social del sujeto. Esto es, el ítem contiene una parte relacionada con un hipotético factor de consumo de sustancias y otra parte referida a otras circunstancias. 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 (modelos FASEM o no estándar) (Bentler, 1993).

Otra ventaja que se ha señalado frecuentemente a favor del análisis confirmatorio es la posibilidad que tiene el investigador para establecer relaciones entre los factores (Bollen, 1989). En el acercamiento tradicional, o ningún factor correlaciona -rotación ortogonal- o todos lo hacen -rotación oblicua-, mientras que en el acercamiento confirmatorio puede establecerse *a priori* un conjunto de condiciones más flexibles en torno a la relación entre los factores; por ejemplo, que dos correlacionen entre sí y otros dos no estén correlacionados.

En cualquier caso, el análisis factorial confirmatorio siempre precisa de la existencia de una teoría articulada que sirva de base para la elaboración de un modelo cuya contrastación empírica se está analizando. Por ello, es preciso que el modelo muestre un buen ajuste a los datos y, además, que los parámetros que componen el modelo muestren la dirección y significación previstas. En el caso específico de la figura 1, esperaríamos que, de ser correcto el modelo de un solo factor global de la autoestima, el modelo presentará un buen de ajuste a los datos y que, además, las saturaciones de

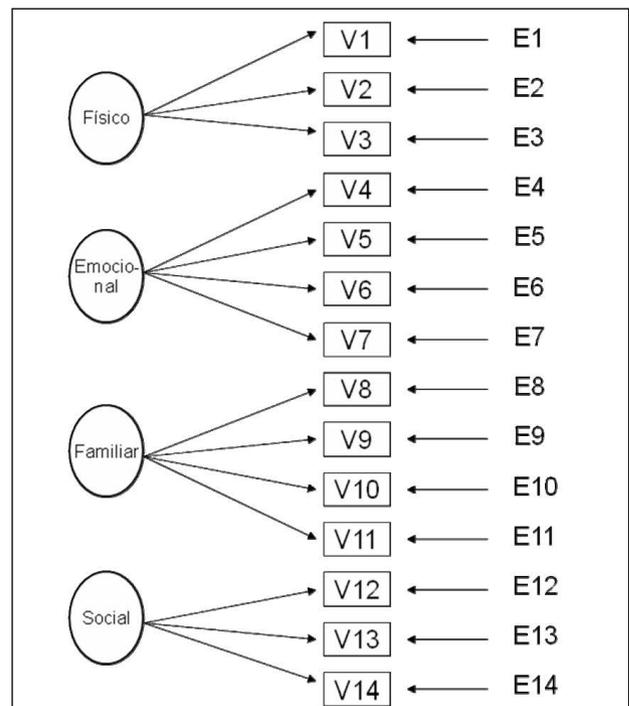
los ítems en el factor fueran todas significativas. Sin embargo, este modelo de un solo factor global de autoestima es uno de los posibles modelos que pueden someterse a contrastación empírica. Es conveniente que el investigador evalúe no sólo un tipo de modelo sino varios que pueden considerarse alternativos a su propuesta teórica, con el fin de conocer hasta qué punto los datos se ajustan mejor a su modelo propuesto que a otro conjunto de modelos alternativos.

Modelos alternativos

La figura 1 representa una conceptualización de la autoestima en términos globales o unidimensionales. Este sería el caso, por ejemplo, del cuestionario de Rosenberg (1965), en el que sólo existe una dimensión global para todos los ítems, si bien existen referencias en la literatura científica sobre su multidimensionalidad (Goldsmith, 1986). Este tipo de concepción unidimensional del self no es la única y otros autores mantienen modelos alternativos en los que se reconoce la complejidad y multidimensionalidad del self (Marsh y Shavelson, 1985). Estas conceptualizaciones alternativas de la estructura de la autoestima se entienden en el Análisis Factorial Confirmatorio como modelos alternativos que habría que estimar y, si es posible, comparar, para tratar de identificar el modelo que mejor describe los datos.

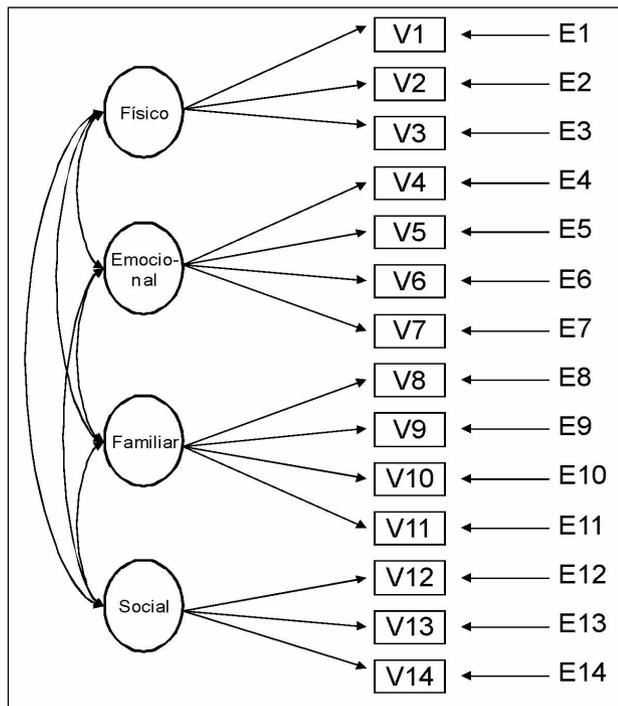
Presento a continuación algunos modelos alternativos que se derivan de las distintas concepciones teóricas de la autoestima.

Figura 2. Modelo 2: 4 dimensiones de la autoestima sin correlacionar



En el modelo 2 la estructura de la autoestima está configurada por la existencia de 4 factores: físico, emocional, social y familiar. La característica fundamental de este modelo es que los factores no muestran ninguna relación entre sí, una circunstancia similar a la rotación ortogonal en el análisis factorial tradicional. Tampoco existe una estructura subyacente a estos factores, como podría ser una autoestima global o factor de segundo orden. Una solución similar al modelo 2 se observa en el modelo 3.

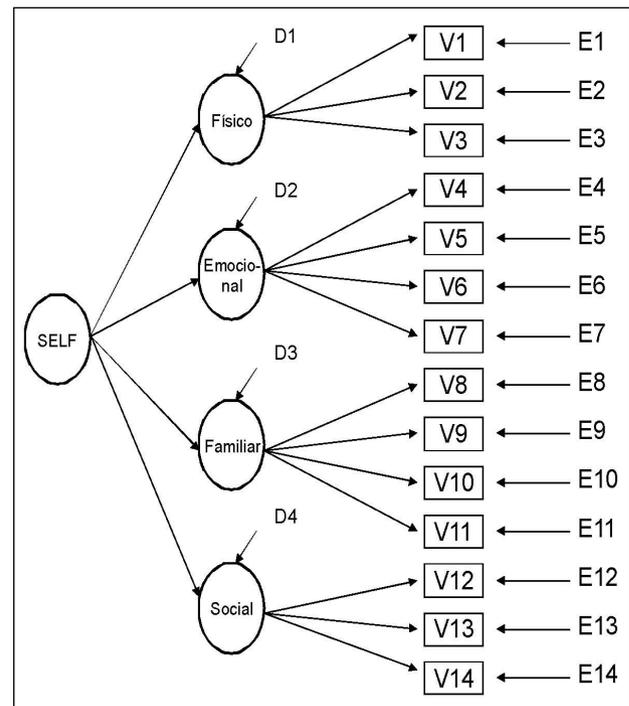
Figura 3. Modelo 3: 4 dimensiones de la autoestima correlacionadas



Existe una estructura factorial con 4 factores que, a diferencia del modelo anterior, se relacionan todos entre sí formando un entramado de interrelaciones. Esta circunstancia es similar a la rotación oblicua en el análisis factorial tradicional.

Finalmente, en el modelo 4 toda la relación observada en las dimensiones de la autoestima del modelo 3 es explicada por la existencia de un factor de segundo orden (SELF). Estadísticamente, los modelos 3 y 4 son muy similares y, en la práctica, es habitual encontrar niveles equivalentes de ajuste en este tipo de modelos con factores correlacionados (modelo 3) y factores de segundo orden (modelo 4). Como ha señalado Bentler (1993), al suponer una estructura factorial con todos los factores relacionados entre sí, lo que se está suponiendo indirectamente es la existencia de un factor latente que explica, a su vez, esas relaciones. Así, el modelo 4 representa una estructura en cuatro dimensiones que a su vez se organizan en una dimensión global o general del self. Cada una de las dimensiones de primer orden del modelo –física, emocional, social y familiar.– está influida por

Figura 4. Modelo 4: Estructura de la autoestima con Factor de segundo orden



el SELF. Sin embargo, las puntuaciones en cada una de estas dimensiones no sólo dependen del factor general SELF sino que están influidas por términos de error o residuales. Así, cada dimensión o factor lleva asociada un error (D) que expresa aquella parte de la varianza de cada factor que no es explicada por el factor general SELF. Una característica importante de este modelo, como ya se ha señalado, es que al tener en cuenta el error asociado a cada indicador, los factores latentes están libres de error de medida, ya que estos errores son parámetros distintos a las variables a las que van asociados. De este modo, el SELF así operativizado es un factor global de autoestima libre de los errores de medida que tienen que ver con los indicadores.

Método

Participantes

Una vez identificado el conjunto de modelos que se van a someter a contrastación empírica, paso a presentar la muestra que se ha utilizado en los análisis. Todos los participantes completaron el CA-14 en dos momentos temporales separados por seis meses. Junto con el instrumento de autoestima, también se recabó información sobre algunas variables como el sexo, la edad, el estado civil, la situación laboral, nivel cultural e ingresos familiares y algunas variables referidas al bienestar bio-psico-social (Gracia, Herrero y Musitu, 2002, para un análisis detallado).

El N final de participantes con datos completos en el pase 1 y el pase 2 fue de $N = 780$ (ver Gracia y Herrero, 2004, para una presentación más detallada de los participantes). El porcentaje de hombres fue ligeramente inferior al de mujeres (46%), con representación en todos los grupos de edad a partir de 18 años (18-24, 22.5%; 25-49, 43.9%; 50-64, 25.2%; > 64, 8.4%). En cuanto al estado civil, la mayoría de los participantes estaba casado (49.1%) o soltero (43.1%). Más de un 54 % de los participantes habían cursado Bachillerato.

Análisis. Algunas precauciones a tener en cuenta

Tamaño de la muestra y bondad de ajuste del modelo

La teoría estadística que sirve de base a los modelos de ecuaciones estructurales con variables latentes, de los que el análisis confirmatorio es un caso particular, es de naturaleza asintótica. Esto quiere decir que las conclusiones que se pueden extraer sobre los datos - con respecto a la distribución de los estadísticos de ajuste y los errores estándar para los parámetros estimados- adquieren cierta confianza conforme N crece sin límite. A pesar de que una muestra muy numerosa permite conclusiones menos inciertas que una muestra pequeña, la verdad es que no hay un límite que indique cuándo una muestra es lo suficientemente grande como para asegurar la confianza en la estimación de los parámetros y el ajuste del modelo. El investigador interesado en la utilización de modelos de ecuaciones estructurales con variables latentes está atrapado en un “doble vínculo” ya que necesita una muestra numerosa para ganar confianza en la estimación de los parámetros del modelo pero, a la vez, la probabilidad de rechazar el modelo en esa muestra es alta - el poder estadístico de la prueba es muy elevado y detectará significación estadística incluso en el caso de que las diferencias sean triviales-. En otros términos, el estadístico resultante de comparar la semejanza entre la estructura de los datos de la muestra, por una parte, y la estructura que propone el modelo, por otra, se distribuye como χ^2 que para grandes muestras tiende a rechazar H_0 (que el modelo es bueno).

Esta circunstancia plantea la necesidad de responder a la pregunta de cuándo el tamaño de la muestra es adecuado y, de ser así, cómo calcular el ajuste del modelo sin depender del tamaño de la muestra. Con respecto a la primera parte de la pregunta, no existe un consenso en los investigadores en cuanto al número de sujetos necesarios para que las estimaciones del análisis confirmatorio sean fiables. Más que un N determinado, lo que sí parece claro es que la fiabilidad del modelo depende mucho de su complejidad y del número de sujetos con que cuenta el investigador para contrastarlo (Jackson, 2003; Kline, 2005; Muthén y Muthén, 2001). A este respecto no existe un acuerdo entre los investigadores sobre cuál es el N adecuado

para estimar un modelo, ya que ello dependerá de la complejidad del modelo y de si se han realizado modificaciones *post-hoc* en el mismo.

Con respecto a la segunda parte de la pregunta, la gran mayoría de los trabajos en esta área de estudio han optado por complementar los índices de ajuste basados en la significación estadística del χ^2 con otro conjunto de indicadores de ajuste de naturaleza variada. Debido a que, como se ha señalado repetidamente, la utilización de grandes muestras aumenta la probabilidad de rechazar la H_0 siendo esta verdadera se han ido proponiendo un conjunto de índices de ajuste que pretenden contestar a la pregunta de hasta qué punto es adecuado un determinado modelo (Bentler, 1990; Bollen y Long, 1993; Hu y Bentler, 1998, 1999; Widaman y Thompson, 2003; Yuan, 2005). La mayoría de estos índices no son estadísticos y, por tanto, los puntos críticos más allá de los cuales un modelo es adecuado, son totalmente arbitrarios y adoptados por consenso en la comunidad científica.

Los paquetes estadísticos en la actualidad proporcionan una gran variedad de índices de ajuste, incluso cuando estos ya no se consideran apropiados en la literatura científica (por ejemplo, NFI del EQS; GFS, AGFS, en LISREL, etc). Esta abundancia de indicadores genera en ocasiones confusión al investigador, máxime cuando alguno de estos índices tienen tendencia a sobrevalorar el ajuste de los modelos, pudiendo llevar a la falsa conclusión de que el modelo es adecuado cuando no lo es. En la práctica, si un modelo presenta un buen ajuste a través del CFI y del RMSEA conjuntamente, es muy poco probable que el modelo no sea adecuado a los datos. Estos índices de ajuste son, por tanto, una buena guía en la búsqueda del modelo que mejor se ajusta a los datos.

El CFI (Comparative Fit Index) fue desarrollado por Bentler (1992) a partir de un índice previo (BFI) que corrige para evitar que tome valores más allá del rango 0-1. El CFI compara el χ^2 de dos modelos: un modelo independiente que mantiene que no existe relación entre las variables del modelo, y el modelo propuesto por el investigador. Esta comparación se corrige por los grados de libertad (gl) de uno y otro modelo. $CFI = ((\chi^2 \text{ Modelo Independiente} - gl) - (\chi^2 \text{ Modelo propuesto} - gl)) / (\chi^2 \text{ Modelo Independiente} - gl)$. Conforme el χ^2 del modelo propuesto disminuye, el numerador y el denominador se igualan, por lo que la situación ideal es que ambos sean equivalentes ($CFI = 1$). Esto es, que el χ^2 del modelo propuesto sea cero. En general, se considera que el CFI debe estar en torno a .95 para considerar que el modelo se ajusta adecuadamente a los datos. Este valor, sin embargo, es relativo ya que, por ejemplo, en modelos de gran complejidad el χ^2 siempre se alejará de cero, lo que hace disminuir el CFI. La interpretación del índice CFI, por tanto, se debe valorar conjuntamente con otros índices y teniendo en cuenta el tipo de modelo que se está analizando.

El RMSEA (error de aproximación) hace referencia

a la cantidad de varianza no explicada por el modelo por grado de libertad. Se considera que $RMSEA \leq .05$ indica un buen ajuste a los datos si, además, el intervalo de confianza al 90% (I.C.) se sitúa entre 0 y .05. Como ya he comentado, en la práctica es poco probable que el investigador encuentre un modelo poco adecuado que cumpla los requisitos: $CFI \geq .95$ y $RMSEA \leq .05$ (90% I.C., .00, .05). Se considera por tanto aconsejable presentar estos dos índices de ajuste junto con el χ^2 del modelo propuesto, sus grados de libertad y la probabilidad asociada.

Estimación del modelo y violación del supuesto de multinormalidad de las variables

El cálculo de los parámetros que componen el modelo puede basarse, a su vez, en distintos métodos de estimación, que consisten, básicamente, en funciones a minimizar. De este modo, si de lo que se trata es de minimizar la suma de los cuadrados estamos hablando del método de Mínimos Cuadrados, mientras que otros métodos, como Máxima Verosimilitud, tratan de minimizar una función algo más compleja (Kaplan, 2000; Olsson, Foss, Troye y Howell, 2000). Existe una extensa literatura sobre la idoneidad de los diferentes métodos de estimación. Una gran parte de los trabajos científicos en este área utilizan la estimación por Máxima Verosimilitud (Bollen, 1989), cuyas propiedades han sido destacadas frecuentemente. La utilización de la estimación por Máxima Verosimilitud (a partir de ahora, ML, del inglés Maximum Likelihood) descansa en un conjunto de condiciones con respecto a la naturaleza de los datos. Entre estas condiciones destaca: a) que se utiliza una muestra suficientemente grande para la estimación de los parámetros; b) que las variables observadas son continuas, y; c) que las variables se distribuyen multinormalmente. Si estas condiciones no se cumplen no hay garantía de que las propiedades de la estimación por ML se mantengan.

El problema de la muestra ya ha sido analizado con anterioridad. Con respecto a la continuidad de las variables, la violación de este supuesto es relativamente asumible y sólo adquiere notoriedad cuando, además, la distribución de las variables no es multinormal (West, Finch y Curran, 1996; Lei y Lomax, 2005). Aunque con la violación del supuesto de multinormalidad, el método ML es relativamente robusto en el cálculo de los parámetros del modelo (coeficientes de regresión o paths entre las variables, etc), no lo es en el cálculo del χ^2 y por tanto en el cálculo de los índices de ajuste, ni en los errores estándar asociados a cada parámetro del modelo (Bentler, 1993; Bollen, 1989; Kaplan, 2000; Marcoulides y Schumacher, 2001). En otras palabras, no es fiable ni en los índices de ajuste del modelo ni en la significación estadística de los parámetros (para cuyo cálculo contribuyen los errores estándar).

La multinormalidad de las variables es, por tanto, un requisito importante para el cálculo de los parámetros y ajuste de un modelo. Sin embargo, es bastante común que el investigador se encuentre con que las variables observadas no sigan una distribución multinormal debido a que esta es una condición relativamente exigente. Como ha señalado Stevens (1992, p. 245), “la normalidad de cada una de las variables por separado es una condición necesaria pero no suficiente para que exista la normalidad multivariada (o multinormalidad)”. Esto es, además de que cada una de las variables medidas muestre unos índices de curtosis y simetría “normales”, existen dos condiciones adicionales que deben cumplir las variables para ser multinormales: a) cualquier combinación lineal de las variables es también normal, y; b) todos los conjuntos y subconjuntos de variables tienen una distribución normal. En la práctica, conforme aumenta el número de variables observadas con que trabaja el investigador la probabilidad de que no exista normalidad multivariada aumenta, ya que también lo hace el número posible de combinaciones entre las variables.

Existen varios procedimientos para paliar en parte este problema. Como han señalado West y colaboradores (1995), todos estos métodos comparten un objetivo común: proporcionar un χ^2 así como estimadores del error asociado a cada parámetro que estén lo más cerca posible de sus valores reales. De lo contrario, tanto el grado de ajuste del modelo (que se basa en el χ^2) como la significación estadística de los parámetros del modelo (que se basan en el error de estimación de cada parámetro) pueden estar sesgados y no ser fiables. Ante esta circunstancia, es prioritaria la contrastación del modelo con procedimientos que aseguren que la desviación de la multinormalidad se ha tenido en cuenta y que tanto el χ^2 como los errores de estimación se aproximan a sus valores reales en la población.

Aunque el estadístico no sigue una distribución χ^2 en condición de no multinormalidad, se puede reescalar o corregir para que se aproxime a ese tipo de distribución. Esta es precisamente la estrategia que han seguido Satorra y Bentler (1990). Básicamente, se calcula una constante que incorpora el grado de curtosis multivariada y se corrige el estadístico (χ^2) por esa constante. La misma estrategia se utiliza para el cálculo de los errores estándar. De este modo, el método de estimación sigue siendo el mismo (ML, pongamos por caso) pero con estimadores robustos que ya tienen en cuenta la desviación de la multinormalidad en las variables. No exige muestra adicional: funciona adecuadamente con un $N=200$ y una ligera desviación de la multinormalidad, y $N=500$ y una desviación sustancial de la multinormalidad. Habitualmente, este índice de ajuste robusto es mayor que el CFI bajo teoría normal (no robusto) ya que utiliza para su cálculo el χ^2 escalado, que suele ser menor que el χ^2 bajo teoría normal. En otras palabras, al tener en cuenta la naturaleza no multinormal de las

variables, el cálculo del χ^2 es más preciso y el índice de ajuste mayor.

Resultados

Una vez presentada la técnica del análisis factorial confirmatorio y analizadas sus ventajas y desventajas, así como la forma elegida para superar esas desventajas, pasamos a continuación a exponer los resultados obtenidos. Debido a que una preocupación inicial en la literatura científica sobre modelos estructurales con variables latentes es precisamente la existencia y tratamiento de la no multinormalidad, se presenta una estimación de la curtosis multivariada a partir del coeficiente de Mardia (1970). Este estimador normalizado se puede utilizar para contrastar la hipótesis nula de que las variables se distribuyen multinormalmente, ya que en muestras suficientemente grandes este estimador se distribuye de forma normal, con puntuaciones extremas positivas indicando curtosis multivariada positiva y puntuaciones extremas negativas indicando curtosis multivariada negativa. En nuestro caso, el estimador es lo suficientemente grande (27.05) como para rechazar la hipótesis de multinormalidad.

En la tabla 1 se presentan los resultados del ajuste de los cuatro modelos analizados teóricamente. En los cuatro análisis se ha utilizado un método de estimación robusto para el cálculo del χ^2 escalado y de los errores estándar, debido a los indicios razonables de no multinormalidad entre las variables.

Los modelos 3 y 4 muestran un ajuste a los datos sensiblemente superior a los dos modelos previos. Estos dos modelos, cabe recordar, introducen la condición de que las dimensiones de la autoestima están interrelacionadas, bien sea permitiendo que correlacionen todas entre sí (modelo 3) o generando un factor general de segundo orden (modelo 4). Si bien el ajuste de estos modelos dista de valores aceptables (CFI robusto = .90), la reducción en el χ^2 en estos modelos es considerable en comparación con los modelos 1 y 2, lo que sugiere que las dimensiones de la autoestima están relacionadas. La inclusión de tres errores correlacionados en el modelo 4 produce una reducción significativa del χ^2 que lleva al modelo a un ajuste adecuado.

Si comparamos exclusivamente los modelos 3 y 4, cabe preguntarse: ¿cuál de los dos modelos es más adecuado? Ante esta pregunta existen dos tipos de respuestas: una respuesta basada en criterios exclusivamente empíricos y una respuesta de tipo conceptual o teórico.

a) Criterios empíricos: Ambos modelos presentan un ajuste similar y puesto que no están anidados, no se pueden comparar estadísticamente. Como ya he avanzado, los modelos con factores correlacionados y los modelos con factores de segundo orden que expliquen esas correlaciones suelen presentar un ajuste similar. Si estuvieran anidados, comprobaríamos que el modelo 4 es preferible por su sencillez: tiene más grados de libertad (2) y estos no le llevan a un incremento significativo del χ^2 (el incremento es = 4.32 que para 2 grados de libertad tiene una probabilidad asociada de $p = .112$). Sin embargo, estos modelos no están anidados y la elección debe basarse en otros criterios.

b) Criterios conceptuales o teóricos: En último término, la elección de un modelo frente a otro debe basarse en criterios teóricos. De no ser así, la configuración de los modelos estaría guiada por la obtención de unos índices de ajuste adecuados que, quizás, fueran asociados a modelos sin ningún contenido teórico. Si bien las cualidades empíricas de ambos modelos pueden considerarse adecuadas, es preciso hacer una reflexión en cuanto a sus implicaciones o cualidades teóricas. En primer lugar, suponer que existe una intercorrelación entre factores sin añadir el “por qué” de esa intercorrelación (modelo 3) conlleva una menor apuesta teórica que afirmar que esos factores se relacionan porque comparten todos ellos la cualidad de representar el SELF (modelo 4). En segundo lugar, el modelo de segundo orden es más sencillo estadísticamente que el de primer orden y, por tanto, más parsimonioso. Además, toda vez que la literatura científica recoge la idea de un self estructurado y jerárquico, el ajuste adecuado a los datos que presenta este modelo de segundo orden supone una confirmación de la existencia de esta estructura. En tercer lugar, como veremos más adelante, las saturaciones de

Tabla 1. Índices de ajuste para los modelos de la estructura factorial del CA-14¹

Modelo	χ^2	df	CFI	RMSEA	Satorra-Bentler χ^2	CFI Robusto	RMSEA Robusto
Modelo 1	950.219	77	.55	.13 (.12, .13)	769.278	.56	.11 (.11, .12)
Modelo 2	479.904	77	.79	.09 (.08, .09)	405.736	.79	.08 (.07, .08)
Modelo 3	265.163	71	.90	.06 (.05, .07)	223.061	.90	.06 (.05, .06)
Modelo 4	268.965	73	.90	.06 (.05, .07)	227.385	.90	.05 (.05, .06)
Modelo 4 Modificado	177.848	70	.95	.05 (.04, .06)	151.480	.95	.04 (.03, .05)
Modelo 4 Modificado Pase 2	196.578	70	.93	.05 (.04, .06)	160.284	.94	.04 (.03, .05)

¹ Se omite la probabilidad asociada a todos los χ^2 de la tabla que es $p < .001$

cada factor en el factor de segundo orden son de moderadas a relativamente elevadas; lo que supone un nuevo indicio de que, efectivamente, esta solución de segundo orden está reflejando una estructura plausible de las respuestas de los sujetos de la muestra al cuestionario.

Valores y significación estadística de los coeficientes de regresión estimados

El cálculo de un modelo implica el cálculo de los parámetros que lo configuran así como su error estándar asociado, lo que posibilita realizar un contraste de la hipótesis de que el valor del parámetro es cero en la población (H_0). El conjunto de parámetros del modelo lo componen tanto las varianzas como los coeficientes que expresan la relación de unas variables con otras (Bentler, 1992, para un análisis más detallado). La versión estandarizada de estos coeficientes equivale a un coeficiente de regresión parcial, en el que se ha extraído el efecto que otras variables puedan tener en esa relación. En el caso del análisis confirmatorio, las relaciones entre las variables o indicadores y los factores latentes tienen una lectura similar a las saturaciones que se obtienen en el análisis factorial exploratorio

Se presentan a continuación los coeficientes relativos a las saturaciones de cada indicador (ítem) en su factor, junto con su significación estadística así como las saturaciones de cada factor en el factor de segundo orden. El modelo postula que el coeficiente de regresión de un ítem es cero para todos los factores excepto para uno, en el que satura. Esta circunstancia marca ya una diferencia con el análisis factorial exploratorio en el que los ítems saturan en todos los factores, si bien el coeficiente de regresión en alguno de ellos es muy pequeño. Los resultados se presentan en la siguiente tabla.

Como se puede apreciar en la tabla 2, los signos de los coeficientes indican que a mayor puntuación en el factor, mayor también es el nivel de autoestima, por lo que es de esperar que todos estos factores saturan positivamente en el factor de segundo orden denominado “self” o “autoestima general”. La última fila de la tabla 2 ilustra la información relativa al modelo de segundo orden. Como se aprecia en esta tabla, las saturaciones en el factor de segundo orden son de moderadas a altas (desde .49 para hasta .81). Además, todos estos coeficientes tienen una probabilidad de ser cero en la población de $p < .001$, lo que lleva a afirmar con un razonable nivel de certeza que la estructura jerárquica del self no sólo se ajusta bien a las respuestas de los sujetos al cuestionario CA-14 sino que, por las magnitudes observadas en las saturaciones en el factor general, la multidimensionalidad de la autoestima permite una importante contribución de cada dominio al self general.

Algunas comprobaciones para verificar si el cálculo del modelo es adecuado

El cálculo de los modelos de ecuaciones estructurales con variables latentes conlleva la estimación de numerosos parámetros y relaciones que conviene examinar en detalle para descubrir posibles anomalías en el comportamiento de los datos que quizás pudieran invalidar los resultados. Si bien los resultados estelares de este tipo de análisis son los índices de ajuste del modelo y el cálculo de las regresiones parciales entre las variables observadas y latentes (McDonald y Ho, 2002; Schreiber, Nora, Stage, Barlow y King, 2006), es conveniente examinar con atención otros parámetros cuyo comportamiento puede apuntar hacia la existencia de problemas en la estimación del modelo. Entre estos parámetros que habitualmente no se analizan en

Tabla 2. Estructura factorial del CA-14. Los coeficientes son estandarizados

	Físico	Emocional	Familiar	Social
Tengo poca resistencia física	-.564			
Tengo una salud excelente	.583			
Tengo partes de mi cuerpo que me gustaría cambiar	-.376			
Me excito con facilidad		-.639		
Soy nervioso/a		-.561		
Soy equilibrado emocionalmente		.441		
Me cuesta controlarme		-.685		
Me siento querido en mi familia			.868	
Me siento feliz en mi familia			.794	
Mis relaciones familiares son insatisfactorias			-.401	
Mi ideas, consejos y opiniones son bien valoradas en mi familia			.401	
Pierdo fácilmente amigos				-.654
En general, no se valora mi amistad				-.586
Mis relaciones sociales son insatisfactorias				-.638
Factor General de Segundo Orden	.604	.527	.491	.811
R ²	.364	.278	.241	.658

los trabajos que utilizan este tipo de técnicas, destacan las varianzas y las covarianzas estandarizadas. Así, por ejemplo, es interesante comprobar que no existen varianzas negativas (casos Heywood) ni que ninguna de las correlaciones entre los parámetros estimados excede a ± 1 . De existir este tipo de circunstancias, los resultados del modelo quedarían invalidados ya que el modelo ofrece una solución que “no es posible” -que una varianza sea negativa o que una correlación exceda ± 1 . Los resultados del análisis, no mostrados en este trabajo, indican que no existe ningún problema adicional en el cálculo del modelo.

Modificación post-hoc del modelo 4

Como ya he mencionado, el modelo 4 incluye en su versión final 3 errores correlacionados cuya estimación permite que el modelo alcance un nivel de ajuste a los datos adecuado. La inclusión de estos parámetros en el modelo se ha realizado *post-hoc*: una vez estimado el modelo, el programa (en este caso EQS) indica los parámetros que podrían liberarse para mejorar el ajuste a través de los índices de modificación del modelo. Esto es, en el modelo inicial estos parámetros tienen un valor de cero pero los resultados sugieren que se deberán estimar libremente ya que ello mejorará el ajuste final del modelo. Una de las críticas más comunes al Análisis Confirmatorio es, precisamente, el hecho de que a menudo los investigadores incluyen nuevos parámetros en el modelo con el mero razonamiento de que mejoran el ajuste, pero sin una justificación teórica sólida. Al final de este proceso, se presentan modelos con una bajísima probabilidad de ser replicados en muestras diferentes a las utilizadas para estimarlos: son trajes a medida de una muestra determinada con muy poca probabilidad de representar a la población. En ocasiones, sin embargo, los índices de modificación permiten identificar algunos aspectos en el modelo que el investigador no haya tenido en cuenta y que pudieran justificarse teóricamente. Sólo en este caso es recomendable utilizarlos para mejorar el ajuste.

En el caso que nos ocupa del modelo 4, se han incluido 3 errores correlacionados. Para analizar su idoneidad seguiré una doble vía: justificar teóricamente la inclusión de estos nuevos parámetros en el modelo y tratar de replicar el modelo modificado con los datos del segundo pase. Con respecto a la primera vía, los errores correlacionados se refieren a los siguientes pares de ítems: (*Me excito con facilidad* y *Soy nerviosa*), (*Mis relaciones familiares son insatisfactorias* y *Mis relaciones sociales son insatisfactorias*), y (*Me siento querido en mi familia* y *Mis relaciones familiares son insatisfactorias*). Una explicación teórica tentativa sugiere que los errores correlacionados entre estos ítems quizás tengan que ver con la dificultad de comprensión de los mismos. Cabe recordar que son

ítems tipo Likert que van de muy en desacuerdo a muy de acuerdo, por lo que alguno de los ítems incurren en una doble negación. Este problema de comprensión también es posible que afecte al ítem *me excito con facilidad*, que quizás fue malinterpretado por algunos participantes. En principio esta explicación basada en la semántica de los ítems es plausible. Complementaria a esta explicación estaría la segunda vía: si los errores correlacionados pertenecen al modelo, también los encontraremos con datos diferentes. En este sentido, el pase 2 puede servir para contrastar esta idea.

Los resultados del modelo 4 modificado para los datos del pase 2 sugieren un ajuste adecuado: (χ^2 , 73 = 196.578, CFI robusto = .94, RMSEA robusto = .04, I.C., .03, .05) muy similares a los ya comentados para los datos del pase 1. Esto constituye una prueba adicional de que el modelo 4 modificado es un buen modelo.

Análisis de la Invarianza Factorial y estabilidad temporal del CA-14

El cálculo del modelo 4 modificado en los datos del pase 2 es una aproximación al estudio de la invarianza factorial: una misma estructura factorial se replica en dos pases. Afortunadamente, el Análisis Confirmatorio permite un estudio mucho más detallado de la invarianza factorial en dos muestras. En este caso, la hipótesis nula (H_0) afirma que el modelo es idéntico para las diferentes muestras y que éste, por tanto, reproduce los datos de cada muestra utilizada en el análisis con un grado razonable de precisión. Conjuntamente, en este tipo de análisis se somete a contrastación estadística la igualdad de los coeficientes en cada una de las muestras o grupos. Por ejemplo, permite someter a contrastación empírica la idea de que las saturaciones de cada ítem sean idénticas en todas las muestras o que cada factor de primer orden sature con la misma magnitud en el factor de segundo orden en todas las muestras. Por último, este tipo de análisis ofrece un χ^2 general que expresa el grado en que el modelo se ajusta simultáneamente a los datos de las diferentes muestras; consecuentemente, es también posible el cálculo de índices de ajuste que utilizan este χ^2 .

De lo que se trata, por tanto, es de estimar el mismo modelo en dos muestras imponiendo a su vez algunas condiciones que debe cumplir este modelo si existe una invarianza factorial. De entre las condiciones que un investigador puede desear contrastar en un análisis confirmatorio en diferentes muestras, he seleccionado:

1. Que las saturaciones en los factores sean iguales en todas las muestras. Si las variables observadas están midiendo los mismos factores en cada muestra, las saturaciones deberían ser las mismas para cada muestra.
2. Que los coeficientes entre los factores sean iguales en todas las muestras. En nuestro caso, que las saturaciones en el factor de segundo orden fueran

equivalentes en las diferentes muestras. Es decir, que la estructura de segundo orden sea igual a través de las muestras.

3. Que las covarianzas (las correlaciones en su versión estandarizada) de los errores asociados a los indicadores sean iguales entre las muestras.

Se ha calculado el modelo 4 modificado en dos muestras simultáneamente (ver tabla 3), estableciendo las siguientes constricciones al modelo:

- a) Las saturaciones de cada ítem en su factor es igual en las dos muestras
- b) Las saturaciones de cada factor en el factor general es igual en las dos muestras
- c) Las correlaciones entre los errores asociados a los indicadores (E) son iguales en las dos muestras.

Relajando esas equivalencias entre el pase 1 y 2 se mejora aún más el modelo tal y como se observa en la reducción del χ^2 . En este punto debo hacer notar que la reducción del χ^2 conlleva una mejora del CFI y RMSEA que no se aprecia en la tabla porque afecta al tercer decimal. Este último modelo puede considerarse como un modelo final de la invarianza factorial del CA-14 a través del tiempo. Analizando en este último modelo la magnitud de los coeficientes sin estandarizar correspondientes a las tres equivalencias no impuestas, se observa que en el pase 2 un ítem presenta una saturación mayor (*Mi ideas, consejos y opiniones son bien valoradas en mi familia*), así como que dos factores de primer orden (familiar y emocional) presentan también en este pase saturaciones mayores. Más allá de estas

Tabla 3. Índices de ajuste para los modelos de la invarianza factorial y estabilidad del CA-14¹

Modelo	χ^2	df	CFI	RMSEA	Satorra-Bentler χ^2	CFI Robusto	RMSEA Robusto
Dos Pases Inicial	2.374.985	341	.65	.09 (.09, .09)	2.012.477	.65	.08 (.08, .08)
E's y D's correlacionadas	731.941	323	.93	.04 (.04, .05)	612.901	.94	.04 (.03, .04)
Saturaciones y correlaciones iguales	669.276	333	.94	.04 (.03, .04)	561.924	.95	.03 (.03, .04)
Modelo Final	652.653	330	.95	.04 (.03, .04)	548.144	.95	.03 (.03, .04)

¹ Se omite la probabilidad asociada a todos los χ^2 de la tabla que es $p < .001$

En la tabla 3 se presentan los resultados de varios modelos estimados para el estudio de la invarianza factorial. El primer modelo no impone ninguna igualdad y deja libre la estimación de todos los parámetros en los dos pases. Es un modelo con un ajuste muy pobre (CFI robusto = .65). ¿Por qué? Básicamente porque no tiene en cuenta que las fuentes de error a las respuestas del cuestionario son las mismas en los dos pases y, por tanto, deberán ser estimadas por el modelo. Esto simplemente quiere decir que se debe estimar como parámetro del modelo las correlaciones entre los errores del mismo ítem a lo largo del tiempo (E1 en el pase 1 y E1 en el pase 2, etc.). Lo mismo se aplica para los residuales de los factores de primer orden (D1 en el primer pase y D1 en el segundo pase). Añadiendo la correlación de los errores asociados a los ítems y a los factores de primer orden, el ajuste mejora sustancialmente (CFI robusto=.94) y puede considerarse un buen modelo. ¿Es esto suficiente? No. Podría añadirse la condición de que puesto que los ítems son los mismos y la estructura también a los seis meses, las saturaciones deberían ser iguales (por ejemplo, del ítem 1 en su factor tanto en el pase 1 como en el pase 2, etc.) tanto para los ítems como para los factores de primer orden.

Por tanto, a este modelo se debe añadir que las saturaciones de los ítems y de los factores de primer orden son iguales en los dos pases. Este último modelo mejora aún más el ajuste (CFI robusto=.95, RMSEA robusto = .03, I.C., .03, .04). Sin embargo, los resultados muestran que la invarianza no es completa, ya que existen tres equivalencias que no se pueden mantener.

pequeñas diferencias, sin embargo, la estructura factorial del CA-14 muestra un alto grado de invarianza en dos muestras. Cuando se repite el análisis de la invarianza en grupos de hombres y mujeres, la estructura del CA-14 es equivalente para ambos, sin excepciones.

Con respecto a la estabilidad de la autoestima (evaluada a partir de su factor general o de segundo orden) el propio modelo de la invarianza factorial permite estimar la regresión entre el factor de segundo orden en pase 1 y pase 2. El coeficiente estimado es de $\beta = .799$, $p < .001$, lo que sugiere que el factor general de autoestima es muy estable a los seis meses en la muestra analizada.

El coeficiente de autorregresión -esto es, la regresión de una variable sobre sí misma en dos momentos temporales- depende tanto de los cambios intraindividuales como interindividuales. Así, este coeficiente puede ser elevado si; a) existe un elevado nivel de cambio intraindividual en gran parte de los individuos de la muestra, b) si existe un elevado cambio intraindividual en pocos individuos de la muestra, y; c) si los cambios intraindividuales son pequeños en relación con la magnitud de las diferencias interindividuales. En otras palabras, la estabilidad refleja una pluralidad de circunstancias, como que todos los individuos cambien en la misma dirección y magnitud (a), que muy pocos individuos cambien (b), o que los cambios registrados en los individuos, si bien importantes, sean mucho menos marcados que las diferencias registradas entre los individuos. Desde este punto de vista, se puede afirmar que el coeficiente de estabilidad sinteti-

za de alguna manera el cambio relativo en las puntuaciones de cada individuo, teniendo además en cuenta las diferencias con los demás individuos; por tanto, la estabilidad no es equivalente a la ausencia de cambio sino, más bien, a la “estabilidad en el cambio”. Queda fuera de este trabajo un análisis más detallado de aquellos sujetos que vieron aumentada o disminuida su autoestima global a lo seis meses.

Referencias

- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238-246.
- Bentler, P. M. (1992). On the fit of models to covariances and methodology to the Bulletin. *Psychological Bulletin*, *101*, 400-404.
- Bentler, P. M. (1993). EQS. Structural Equations Program Manual. Los Angeles: BMDP Statistical Software Inc.
- Bentler, P. M. (2000). Rites, wrongs, and gold in model testing. *Structural Equation Modeling*, *7*, 82-91.
- Bentler, P. M. (2007). Can scientifically useful hypotheses be tested with correlations? *American Psychologist*, *62*, 772-782.
- Bollen, K. A. y Long, J. S. (1993). *Testing structural equation models*. Newbury Park, CA: Sage.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley y Sons.
- Bollen, K. A. (2000). Modeling strategies: In search of the holy grail. *Structural Equation Modeling*, *7*, 74-81.
- 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.
- Goldsmith, R. E. (1986). Dimensionality of the Rosenberg Self-Esteem Scale. *Journal of Social Behavior and Personality*, *1*, 253-264.
- Gracia, E. y Herrero, J. (2006). La comunidad como fuente de apoyo social: evaluación e implicaciones en los ámbitos individual y comunitario. *Revista Latinoamericana de Psicología*, *38*, 327-342.
- Gracia, E. y Herrero, J. (2004). Determinants of social integration in the community: An exploratory analysis of personal, interpersonal and situational variables. *Journal of Community and Applied Social Psychology*, *14*, 1-15.
- Gracia, E., Herrero, J., y Musitu, G. (2002). *Evaluación de recursos y estresores psicosociales en la comunidad*. Madrid: Síntesis.
- Herrero, J. y Gracia, E. (2007). Measuring perceived community support: Factorial structure, longitudinal invariance and predictive validity of the PCSQ (Perceived Community Support Questionnaire). *Journal of Community Psychology*, *35*, 197-217.
- Herrero, J. y Gracia, E. (2004). Predicting social integration in the community among college students. *Journal of Community Psychology*, *32*, 707-720.
- Hu, L. y Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424-453.
- Hu, L., y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.
- Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N:q hypothesis. *Structural Equation Modeling*, *10*, 128-141.
- Kaplan, D. (2000). *Structural equation modeling: Foundations and extensions*. Thousand Oaks, CA: Sage.
- Kline, R. (2005). *Principles and practices of structural equation modeling* (2nd ed.). New York: Guilford.
- Lei, M., y Lomax, R. G. (2005). The effect of varying degrees of nonnormality in structural equation modeling. *Structural Equation Modeling*, *12*, 1-27.
- Marcoulides, G. A. y Schumacker, R. E. (2001). *New developments and techniques in structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519-530.
- Marsh, H. W., y Shavelson, R. J. (1985). Self-concept: its multifaceted, hierarchical structure. *Educational Psychologist*, *20*, 107-125.
- McDonald, R. P., y Ho, M. (2002). Principles and practice in reporting structural equation analysis. *Psychological Methods*, *7*, 64-82.
- Muthén, L. K. y Muthén, B. O. (2001). *Mplus User's Guide*. Los Angeles: Muthén y Muthén.
- Olsson, U. H., Foss, T., Troye, S. V., Howell, R. D. (2000). The performance of ML, GLS, and WLS estimation in structural equation modeling under conditions of misspecification and nonnormality. *Structural Equation Modeling*, *7*, 557-595.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Satorra, A. y Bentler, P. M. (1990). Model conditions for asymptotic robustness in the analysis of linear relation. *Computational Statistics y Data Analysis*, *10*, 235-249.
- 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.
- Stevens, J. (1992). *Applied multivariate statistics for the social sciences*. Hillsdale, NJ: Earlbaum.
- Tomarken, A. J., y Waller, N. G. (2003). Potential problems with “well fitting” models. *Journal of Abnormal Psychology*, *112*, 578-598.
- West, S. G., Finch, J. F. y Curran, P. J. (1995). Struc-

tural equations models with non-normal variables. En R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications*. Thousand Oaks, CA: Sage

Widaman, K. F. y Thompson, J. S. (2003). On specifying

the null model for incremental fit indices in structural equation modeling. *Psychological Methods*, 8, 16-37.

Yuan, K. H. (2005). Fit indices versus test statistics. *Multivariate Behavioral Research*, 40, 115-148.

Manuscrito Recibido: 19/04/2010

Revisión Recibida: 17/09/2010

Manuscrito Aceptado: 29/09/2010

Apéndice. Medias, desviaciones típicas y correlaciones de los ítems del CA-14 en el primer pase

Ítem	Media	D.T.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	2.56	1.01	-													
2	3.23	.89	-.35	-												
3	3.04	1.17	.18	-.20	-											
4	3.15	.98	.07	-.02	.10	-										
5	3.42	1.09	.12	-.13	.16	.47	-									
6	3.39	.87	-.09	.20	-.21	-.22	-.20	-								
7	2.65	.95	.18	-.15	.17	.44	.33	-.33	-							
8	4.05	.77	-.05	.19	-.14	-.06	-.07	.27	-.14	-						
9	4.16	.79	-.07	.14	-.12	-.02	-.05	.25	-.11	.70	-					
10	2.21	1.04	.03	-.06	.05	.08	.10	-.08	.12	-.31	-.35	-				
11	3.54	.86	-.11	.14	-.08	.01	-.08	.22	-.14	.41	.33	-.20	-			
12	1.97	.78	.20	-.14	.12	.12	.09	.16	.24	-.24	-.23	.18	-.17	-		
13	2.28	.85	.17	-.14	.13	.08	.10	-.10	.21	-.18	-.13	.20	-.16	.40	-	
14	2.28	.84	.19	-.14	.18	.10	.13	-.21	.27	-.25	-.18	.29	-.19	.41	.37	-