Modelos complementarios al Análisis Factorial en la construcción de escalas ordinales: un ejemplo aplicado a la medida del Clima Social Aula

Models to Complement Factor Analysis in the construction of ordinal scales: One customised example applied to Classroom Social Climate

Emelina López González

Universidad de Málaga. Facultad de Ciencias de Educación y Psicología. Departamento de Métodos de Investigación e Innovación Educativa. Málaga, España.

Amparo Pérez Carbonell

Genoveva Ramos Santana

Universidad de Valencia. Facultad de Filosofía y Ciencias de la Educación. Departamento de Métodos de Investigación y Diagnóstico en Educación (MIDE). Valencia, España.

Resumen

El Análisis Factorial (AF) de componentes principales es uno de los modelos estadísticos de mayor aplicación en investigaciones educativas para estudiar las propiedades de una escala. Su generalizado uso se extiende también a escalas no métricas (categóricas u ordinales) y, sin embargo, las propiedades de estas escalas no cumplen con los requisitos del Análisis Factorial (tales como normalidad de las distribuciones o linealidad de las relaciones entre las variables). Son varios los problemas que se generan y algunos los procedimientos destinados a solventar, en mayor o menor medida, esta falta de adecuación del modelo del AF a datos no métricos. Aquí se describen brevemente algunos de los más relevantes: el Análisis Paralelo o el Análisis Factorial desde la Teoría de Respuesta al Ítem. Igualmente, se propone como una posible solución el empleo conjunto de modelos multivariantes de interdependencia de reducción de la dimensionalidad, tales como el Análisis

Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

de Componentes Principales No Lineal (ACPNL), el Escalamiento Multidimensional (EMD) no métrico y el Análisis Cluster (AC), modelos que se adaptan bien a escalas ordinales. En esta línea se muestra un uso combinado y complementario de estos dos últimos, el EMD no métrico y AC, junto con el AF de componentes principales, en el estudio de una escala Likert sobre Clima Social del Aula elaborada en una investigación previa. Como resultado, el ajuste de los modelos a los datos mejora notablemente y la solución de las dimensiones internas es más parsimoniosa y coherente con la concepción sustantiva que inspiró la escala.

Palabras clave: Análisis Factorial, Escalamiento Multidimensional no métrico, Análisis Cluster, escala ordinal, escala Likert, clima social aula.

Abstract

It is well known that factor analysis (FA) by principal components is one of the statistical models more frequently used in educational research to study the properties of scales of measurement. The generalised use of FA extends also to non-metric scales (such as categorical and ordinal scales), yet the properties of such scales do not meet the requirements of factor analysis (such as normal distributions and linear relationships between variables). Several problems are thus generated, but there are a number of techniques for solving to some extent the imperfect application of the FA model to non-metric data. Some of the most important models are briefly described: parallel analysis or factor analysis using item response theory. The joint use of multivariate models of interdependence for dimensionality reduction, such as non-linear principal component analysis (NLPCA), non-metric multidimensional scaling (NM-MDS) and cluster analysis (CA), is proposed as a possible solution, because all these models can be considered appropriate for use with ordinal scales. An example of the combined, complementary use of NM-MDS and CA is shown, together with FA by principal components, to study a Likert scale concerning classroom social climate that was prepared in a prior research project. The results show that the models are significantly more fit for the data, and the internal dimension solution is more parsimonious and consistent with the underlying idea that lay at the origin of the scale.

Key words: Factor analysis, non-metric multidimensional scaling, cluster analysis, ordinal scale, Likert scale, classroom social climate.

Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

Introducción

Es habitual en la elaboración de escalas para medir constructos educativos y psicológicos, el empleo de diversas estrategias estadísticas con el fin de estudiar sus propiedades métricas. Entre las posibles opciones destaca el Análisis Factorial (AF) de componentes principales como una de las técnicas analíticas más utilizada. Sin embargo, es también frecuente que dichas escalas posean una métrica más próxima a niveles de medida categóricos, cualitativos, discretos, nominales, ordinales, todos ellos considerados no métricos. No debe olvidarse que el Análisis Factorial requiere de ciertas condiciones que se ven afectadas por la medida de las variables que se emplean, lo que hace que, aun siendo muy generalizado su empleo, los datos no cumplan los supuestos requeridos. Su extendido uso puede justificarse, entre otras razones, porque pasa por ser una técnica bien competitiva, permitiendo llegar a conclusiones que dan explicación a aspectos sustantivos esenciales de la investigación que se trate, lo que no se consigue fácilmente con otras opciones analíticas sustitutivas. Efectivamente, el hecho de que existan numerosas variables, ítems, elementos, y pueda configurarse una estructura de menores dimensiones contemplando sus relaciones internas, permite detectar si esa configuración subyacente se corresponde con el diseño teórico sustantivo que inspiró la elaboración del instrumento.

La preocupación por el buen uso del AF con datos categóricos no ha pasado desapercibida, como puede verse en Bartholomew (2007, p. 17) y Bartholomew, Steele, Moustaki y Galbrain (2002, cap. 8). Aquí se comentan algunos procedimientos que solventan, en cierta medida, aspectos inadecuados del empleo del AF que utiliza el método de extracción de componentes principales en el estudio de escalas ordinales. Estos procedimientos apuntan, fundamentalmente, en tres direcciones: (a) centrarse en la métrica de la escala realizando pequeñas transformaciones que persigan la depuración de la medida para alcanzar un nivel de intervalo o razón (Schriesheim y Castro, 1996 y Cañadas y Sánchez, 1998); (b) trabajar con distintos desarrollos del AF desde el contexto de la Teoría de Respuesta al Ítem (expuestos con detalle en la obra editada por Thissen y Wainer, 2001), o desde los modelos de ecuaciones estructurales (Muthén, 1984), y (c) emplear de forma complementaria otros modelos multivariantes de interdependencia que, al igual que el AF, busquen una estructura interna en los datos. Hay también adaptaciones del AF a datos categóricos que usan estimaciones bayesianas, y otras referidas a modelos de respuesta de elecciones múltiples (Shigemasu, 1990).

Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

De las opciones señaladas, la (c) es la que aquí exponemos con más detalle: aplicamos un modelo de AF de componentes principales con rotación oblicua a una escala de categorías ordinal –una escala Likert– y, a partir de los resultados, la solución factorial se comenta y amplía con otras estrategias.

Conviene aclarar que hablamos de modelos multivariantes de interdependencia porque analizan todas las variables en conjunto sin mediar funciones de dependencia entre ellas (seguimos la clasificación de Hair, Anderson, Tatham y Black, 1999, p. 15). Son modelos que, lejos de suponer un panorama de alternativas estadísticas al AF entre las que el investigador se vea obligado a optar, pueden usarse de forma complementaria, de manera que la interpretación y configuración dimensional del Análisis Factorial, entre otros resultados, se ve mejorada. Es más, son modelos diseñados para trabajar con escalas no métricas que no reflejan cantidades relativas o de grado, lo que viene a resolver el inconveniente de detectar estructuras subyacentes con valores ordinales. Nos estamos refiriendo a modelos como el Escalamiento Multidimensional (EMD) no métrico, el Análisis de Conglomerados (AC) jerárquicos o el Análisis de Componentes Principales No Lineal (ACPNL).

Problemas asociados al uso del Análisis Factorial y procedimientos propuestos

El esfuerzo por encontrar una estructura subyacente en un conjunto total de estímulos, como puede suponer una escala, es un proceso ineludible en el estudio de las propiedades métricas de instrumentos que se elaboran. A ello debe añadirse la gran profusión en la investigación educativa, en psicología aplicada, en estudios de opinión, de marketing, etc., del empleo de escalas de categorías en las que el sujeto responde a los ítems con un conjunto específico de cuantificadores lingüísticos (Cañadas y Sánchez, 1998), en su mayoría de cantidad (mucho, bastante, poco o nada), como sucede en una escala Likert. Así como estas escalas aportan numerosas ventajas (son fáciles de desarrollar para el investigador, sencillas de comprender, de rápida respuesta) implican, no obstante, algunos problemas que difícilmente pueden eludirse.

La cuestión relativa a los niveles de medida que alcanzan las escalas (como máximo ordinal) no permite asumir que, por el hecho de estar ordenados sus

Revista de Educación, 354. Enero-Abril 2011, pp. 369-397
Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

valores o cuantificadores, lo hagan con igual intensidad como si de una escala de intervalo se tratase, adquiriendo las propiedades métricas de esta última. Este argumento no tiene suficiente base matemática. Para empezar, los datos no pueden ser sumados o promediados, como exige cualquier sencillo cálculo de varianzas o correlaciones lineales, resultando muy comprometido el uso de análisis paramétricos o de cualquier prueba que requiera supuestos, como es el caso del AF de componentes principales (Bartholomew, 2007, p. 16). En el contexto de investigaciones dirigidas a resolver esta cuestión, una solución ha sido optar por desarrollar «nuevos conjuntos de cuantificadores» de las respuestas que sí alcancen el nivel de medida de intervalo. Respecto a esta idea cabe mencionar los trabajos de Schriesheim y Castro (1996) y Cañadas y Sánchez (1998).

Un segundo tema delicado es la correlación entre los ítems con distinta métrica, o con la misma pero ordinal. Los ítems pueden asociarse por su similitud sustantiva (por el contenido), pero también por la similitud de sus distribuciones estadísticas (en este sentido es más fácil que se asocien elementos con distribuciones parecidas que con distribuciones distintas [Bernstein, 1988, p. 398]). De otro lado, los ítems más fáciles de responder o más comprensibles, por ejemplo, suelen obtener una mayor frecuencia de respuesta y por ello agruparse en factores distintos a los elementos de menor frecuencia de respuesta, aunque sustantivamente ambos formen parte de una única dimensión. Puede haber, por tanto, ítems aparentemente multidimensionales asociados a ítems con distribuciones parecidas, cuando en realidad no lo son (Nunnally y Bernstein, 1994, p. 643; Bernstein y Teng, 1989). Hacer interpretaciones sustantivas o conceptuales sobre la naturaleza de estos factores resultantes es seguro que lleva a error.

Existen diversos procedimientos destinados a solventar los problemas de asociaciones de ítems ordinales. Tal es el caso de la propuesta de Muthén (1984) y Muthén y Kaplan (1992) que combina el «análisis factorial con un modelo de ecuación estructural para datos categóricos», empleando un ajuste de mínimos cuadrados generalizado para escalas dicotómicas, o de máxima verosimilitud y mínimos cuadrados ponderados para escalas ordinales (Flora y Curran, 2004). De manera sencilla puede describirse del siguiente modo.

Se tienen en cuenta los dos distintos niveles de variables que acontecen en el AF: los factores o variables latentes (en cuanto a que no son observables) y los indicadores observables de estos factores (los ítems). Se trataría de considerar un nivel intermedio entre ambos ocupado por las «variables latentes de respuesta» (una para cada cuantificador o categoría de respuesta) y que son cortadas en los «puntos umbral» correspondientes a los juicios observados de la escala Likert (mucho, bastante, poco o nada, por ejemplo). La peculiaridad de estas variables latentes de respuesta es que son continuas y han sido dicotomizadas (si hay dos categorías) o politomizadas (más de dos categorías) en determinados puntos umbral, y para analizar la relación entre ellas se emplea la correlación tetracórica o policórica, respectivamente, en lugar de la correlación de Pearson (que es claramente inadecuada). Por tanto, hay dos niveles de abstracción relacionados: por un lado ítems observados ordinales, y por otro variables de respuesta no observables; ambos niveles de variables se asocian por medio de correlaciones policóricas. El AF puede, entonces, aplicarse directamente porque se está trabajando con indicadores continuos aunque no observables, siendo estas variables de respuesta latentes indicadores de los factores. Así, los factores que resultan informan de las relaciones entre las variables de respuesta latentes, más que de las relaciones entre los ítems ordinales que se observaron directamente.

Un enfoque que resuelve satisfactoriamente los problemas asociados a las matrices de correlaciones entre ítems ordinales y que utiliza toda la información contenida en el patrón de frecuencias de las categorías de respuesta es el «análisis factorial de información completa», análisis relacionado con la Teoría de Respuesta al Ítem (TRI). Para datos dicotómicos, este enfoque se encuentra implementado en Bock, Gibbons y Muraki, (1988); su extensión a respuestas politómicas está en Muraki y Carlson (1995) y Swygert, McLeod y Thissen, (2001). Uno de los más recientes trabajos al respecto es el de Gibbons y cols. (2007). La ventaja que aporta radica en la misma concepción de la TRI. La TRI supone que una dimensión subyacente genera un conjunto de respuestas observadas en los ítems ordinales. La probabilidad de responder a un determinado ítem es una función de dos componentes (los parámetros del ítem y los parámetros para el sujeto que responde) y refleja dónde está localizada esta función en el factor subyacente. Cada ítem de la escala se puede representar gráficamente con una función no lineal (por lo general logística) que muestra las propiedades del ítem. La dimensión latente a un conjunto de ítems se llama «habilidad» (un supuesto en los modelos TRI es que los ítems de la escala miden sólo una habilidad –supuesto de unidimensionalidad–). Tanto el AF como la TRI dan información sobre los parámetros de los ítems y de los sujetos. A pesar del supuesto de unidimensionalidad de los ítems, la escala puede descomponerse en escalas unidimensionales más pequeñas antes del análisis, o bien los ítems pueden reagruparse es *testlet*s (Swygert y cols. 2001), que son grupos de ítems que se suponen unidimensionales.

Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

A los problemas señalados sobre empleo del AF con escalas ordinales, hay que añadir los inconvenientes propios del Análisis Factorial en cualquier caso, como la determinación del número óptimo de factores y la trascendencia de las observaciones ausentes (*missing values*). No es extraño que en el AF habitual el número de factores que se obtenga esté sobrestimado, siendo mayor que el que debiera corresponder según argumentos sustantivos. Incluso cuando se realiza un análisis exploratorio y luego un posterior factorial confirmatorio, es muy posible que la estructura obtenida por el primero sea rechazada por el segundo (Ferrando y Lorenzo, 2000). Franklin, Gibson, Robertson, Pohlmann y Fralish (1995) y Buja y Eyuboglu (1993) señalan el «Análisis Paralelo» (AP) como un procedimiento bastante eficaz para indicar el número adecuado de factores significativos, así como para determinar qué ítems deben conformarlo (en O'Connord, 2000 se muestran sencillos programas para poder ejecutar este análisis con SPSS y SAS). El procedimiento puede resumirse del siguiente modo.

Se genera por ordenador una matriz de valores aleatorios y normales de igual dimensión que la matriz de respuestas emitidas por los sujetos («p» ítems y «n» sujetos). Por medio de un AF de componentes principales, y empleando una matriz de correlaciones policóricas con los ítems ordinales observados, se obtienen los autovalores. Antes de cualquier rotación estos autovalores se comparan con los autovalores del AP, que son tomados como «puntos umbrales». Todos los componentes con autovalor superior al punto umbral son considerados espurios. El resultado es que «el AP reduce la interpretación subjetiva de métodos supuestamente objetivos» eliminando la sobre extracción que suele suceder en los factoriales de componentes principales, en general, y específicamente en aquellos que se realizan con variables ordinales (Franklin y cols., 1995).

Respecto a las no respuestas de los sujetos o valores ausentes, en el Análisis Factorial cobran especial protagonismo en el momento de generase la matriz de correlaciones. Es fácil observar cómo el tamaño muestral con el que se comienza a construir la matriz de entrada queda considerablemente reducido, ya que cuando un sujeto no responde a un ítem es eliminado de cualquiera de los emparejamientos entre variables. Existen varios procedimientos para tratar los datos *missing* (pueden consultarse en Little y Rubin, 2002 y Schafer, 1997). Para el caso que nos ocupa, este problema no tiene consecuencias tan trascendentes en los modelos multivariantes de interdependencia complementarios al AF que se proponen en el siguiente apartado.

El uso de métodos multivariantes de interdependencia como alternativa

Volviendo al interés por encontrar la estructura interna de una escala y a la reducción de ítems a un menor número de factores o componentes, efectivamente, o se «arregla» la métrica de la escala o se trabaja con modelos que se ajusten mejor a la medida empleada. Aunque, como hemos ido refiriendo, son diversos los procedimientos para solventar algunos errores, desde nuestro punto de vista en el ámbito de los estudios en educación no hay una alternativa al Análisis Factorial que satisfaga suficientemente, si nos atenemos al escaso empleo por parte de investigadores y profesionales de estos procedimientos. No debe perderse de vista que, si bien durante la primera mitad del siglo XX el AF se desarrolló fundamentalmente a partir del trabajo de psicólogos, en la segunda mitad de siglo le correspondió el dominio de su implementación a estadísticos, quienes avanzaron en los métodos computacionales asociados (Jöreskog, 2007, p. 48). Es evidente que las extensiones y generalizaciones del AF, en muchos casos, no han estado exentas de cierta complejidad matemática. En esta línea, una mejora que sin duda está al alcance de los profesionales en educación es el uso de modelos multivariantes de interdependencia que comparten con el AF la intención de reducir las variables originales, y que probablemente sean más habituales en la literatura sobre investigaciones educativas.

El «Escalamiento Multidimensional no métrico» (EMD) es una alternativa interesante que, igualmente, se encarga de desvelar la estructura oculta de los datos, bien en forma de agrupaciones significativas de las variables, bien en relación a las dimensiones que se emplean, tomadas éstas como referentes para interpretar la solución aportada. Estas similitudes «percibidas» como proximidades o distancias se trasladan a un espacio de pocas dimensiones (habitualmente dos o tres) proporcionando una configuración de coordenadas. Como resultado se obtienen representaciones gráficas que permiten visualizar los posicionamientos de los puntos en un espacio geométrico, o lo que es lo mismo, las distancias entre los ítems asociados a los factores de la escala. De esta forma se detectan las relaciones y las diferencias existentes entre los ítems, y el modo en que pueden ser agrupados.

Es necesario aclarar que, si bien el EMD y AF de componentes principales presentan semejanzas en su interés por determinar una representación geométrica en el espacio con dimensión mínima basándose en los mismos teoremas (teorema de la descomposición espectral –ver Rivas y Martínez Arias, 1991–), y que incluso el procedimiento de cálculo en ambos tiene algoritmos equivalentes en

Revista de Educación, 354. Enero-Abril 2011, pp. 369-397
Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

ciertas etapas (el EMD utiliza también un análisis de componentes), no es menos cierto que las diferencias entre los dos modelos son lo suficientemente importantes como para no aceptar que exista una identidad fundamental entre ambos, y no deben considerarse, por tanto, uno alternativo del otro, o viceversa. Esto redunda en nuestro interés por proponer un empleo conjunto y complementario, y no alternativo o sustitutivo.

Dentro de los distintos tipos de escalamientos multidimensionales, el que aquí interesa es el EMD no métrico (Shepard, 1962 y Kruskal, 1964a, 1964b) ya que, además de suponer una solución satisfactoria al problema de las observaciones ausentes mencionado arriba (Arce y Andrade, 2000), se adapta bien a escalas ordinales y de pequeño recorrido (escala Likert de 1 a 4, como en el ejemplo que exponemos más adelante). Este escalamiento no métrico plantea una relación de tipo monotónica creciente entre las variables de entrada y las derivadas (las distancias), no lineal, que se ajusta muy bien a variables ordinales (Real y Varela, 2003, p. 472). Respecto a las opciones para la construcción de las distancias de partida, es la distancia euclídea la que mejor se adapta a la ordenación secuencial que admite la escala Likert, aunque esta medida sea también habitual con escalas de intervalo, como apuntan Picón, Varela y Real (2003, p. 426).

El «Análisis Cluster» (AC) es otro modelo de interdependencia que clasifica objetos, en este caso ítems, en grupos. Es también un modelo de reducción de la dimensionalidad por dos razones (Cook y Swayne, 2007, p. 104): (a) los ítems quedan resumidos en la descripción de cada grupo, cluster o conglomerado, y (b) cada cluster está compuesto por ítems relativamente homogéneos y puede ser analizado de forma separada a los demás clusters. Ahora bien, el resultado no es una representación geométrica de los ítems en un espacio de menores dimensiones. Para explorar la dimensionalidad en el espacio lo adecuado es el EMD no métrico, y la explicación de las agrupaciones que se construyen y en qué momento se realizan corresponde al Análisis Cluster (Timm, 2004, p. 516). Por tanto, ambos modelos, escalamiento y cluster, son soluciones que se completan y matizan una a otra.

El Análisis Cluster para datos ordinales usa como medida de proximidad la distancia euclídea, igual que el EMD no métrico. A partir de las matrices de distancias entre parejas de ítems se construyen conglomerados jerárquicos adaptados a la escala ordinal (escala Likert), y con un tipo de agrupación aglomerativa o ascendente, iniciándose el análisis con tantos grupos como ítems tenga la escala. A partir de estas unidades se van formando conjuntos de manera ascendente, contemplando cada vez más ítems en los sucesivos grupos. Al final del proceso todos los ítems terminan contenidos en un mismo conglomerado. Este proceso queda bien reflejado en los dendrogramas que resultan.

El «Análisis de Componentes Principales No Lineal» (ACPNL) trabaja con similares objetivos que el Análisis Factorial de componentes principales lineal. La peculiaridad de la estrategia no lineal es que se adapta a variables medidas en escala ordinal, donde cabe esperar que las relaciones entre ellas no sean exactamente lineales. En este análisis se categorizan las respuestas a los ítems con el mismo «orden» original de las variables observadas, por lo que no se produce ninguna pérdida de información, y tampoco se necesita encajar la medida ordinal a algunos de los supuestos paramétricos de las estrategias de análisis más tradiciones, como la linealidad en el AF. De esta forma, el análisis de componentes principales no lineal se adapta en mayor medida a las «verdaderas relaciones» que mantienen los ítems dentro de la escala, superando así la mayor parte de las dificultades que plantea el análisis de la dimensionalidad con datos ordinales, y constituyendo, a nuestro juicio, la alternativa preferible al AF de componentes principales lineal. Esta característica permite establecer un matiz entre las tres opciones multivariantes de interdependencia señaladas. Mientras que el EMD no métrico y el AC jerárquico son alternativas complementarias, es posible considerar el ACPNL como una alternativa sustitutiva.

En el presente artículo hemos optado por presentar la opción (c) sobre estrategias multivariantes complementarias al AF, mencionada arriba en la introducción. Por la importancia y envergadura que supone el empleo del ACPNL como estrategia sustitutiva al AF de componentes principales lineal, dejamos para un trabajo posterior su ilustración. No obstante, para un estudio pormenorizado del ACPNL recomendamos el trabajo de Bolton, Hand y Webb (2003) donde, además, se aplican ambas aproximaciones del Análisis Factorial, lineal y no lineal, y se discuten sus ventajas e inconvenientes.

Metodología

Instrumento

En este estudio se trabaja con una escala diseñada para evaluar la variable de contexto Clima Social Aula, una de las variables relevantes en el desarrollo del

 Revista de Educación, 354. Enero-Abril 2011, pp. 369-397
Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

aprendizaje de los estudiantes (Fernández y Asensio, 1993 y Cid, 2004) y base también para el desarrollo personal del alumno. La elaboración de este instrumento se realizó a través de dos proyectos de investigación: «Evaluación del Clima Social Aula en educación secundaria» (Proyecto Precompetitivo de la Universidad de Valencia, 2005) y «Análisis de Variables de contexto: Diseño de cuestionario de contexto para la evaluación de sistemas educativos» (Proyecto I+D+I, 2005). Como resultado de estos proyectos, se analizaron previamente las dimensiones y reactivos utilizados en otras escalas relacionadas con la variable en cuestión; se definió el constructo en colaboración con un comité de expertos; se diseñó una escala base de Clima Social del Aula y se empleó dicha escala en un estudio piloto. Estos procesos se describen en el reciente trabajo de Pérez-Carbonell, Ramos y López-González (en prensa).

Objetivo

El objetivo aquí es mostrar las ventajas del empleo conjunto del EMD y del AC con el AF de componentes principales para valorar el funcionamiento de la escala sobre Clima Social del Aula obtenida en las mencionadas investigaciones. Más específicamente, a partir de la estructura de dimensiones encontrada con el AF realizado en el estudio piloto, usamos ahora estas otras alternativas complementarias, empleando, por ejemplo, los mapas perceptuales de un EMD, o una distinta configuración de conglomerados resultante de un AC. Como consecuencia, el ajuste de los modelos mejora, lo que permite una interpretación sustantiva más confiable de las dimensiones. Otra de las ventajas es ayudar en el proceso de validez de constructo respecto a los aspectos considerados por los Comités de Expertos como pertinentes y adecuados en la conformación del constructo (lo que en dichos trabajos pasó a denominarse como «Variables de Clima Social Aula», tal y como se describe en Pérez-Carbonell y cols., en prensa).

Muestra y variables

Se trabajó en 14 centros educativos de la Comunidad Valenciana, en el nivel de 4º de Secundaria, con una muestra de 407 alumnos seleccionados con un muestreo no probabilístico.

El cuestionario piloto se compuso de 40 ítems. A partir del AF de componentes principales de rotación oblicua, los elementos se agruparon en diez dimensiones del constructo Clima Social Aula atendiendo a la interpretación sustantiva decidida en las fases anteriores de la investigación. Los factores y la distribución de los ítems se presentan en la Tabla I.

TABLA I. Configuración de factores según el Análisis Factorial

N.º ítem	FACTOR 1: Relación, interés y comunicación	Factor
12	Los profesores se interesan personalmente por cada uno de nosotros.	1
13	Los profesores escuchan a los alumnos sin interrumpir.	1
14	Los profesores muestran respeto por nuestros sentimientos.	1; 8
18	Las relaciones entre nosotros y los profesores son agradables.	1
22	La relación entre los profesores y los alumnos es cordial.	1
30	En esta clase los alumnos tenemos muy buena comunicación con los profesores.	1
32	La mayor parte de los profesores nos animan a hablar en nuestro grupo de clase.	1
47	Hay una buena comunicación entre nuestros profesores.	1
48	Los profesores se sienten orgullosos de esta clase.	1; 9
	FACTOR 2: Cohesión y satisfacción del grupo	
19	Los alumnos estamos contentos con el grupo clase.	2
20	Los alumnos nos sentimos orgullosos de esta clase.	2
25	El aula es un lugar donde me siento solo.	2; 4
26	En esta clase los alumnos nos llevamos muy bien.	2
31	En esta clase los alumnos tenemos muy buena comunicación entre nosotros.	2
35	Los alumnos colaboramos muy bien entre nosotros.	2; 5; 7
44	Creo que mi clase es un lugar agradable (me gusta estar en mi clase).	2
	FACTOR 3: Competitividad	
23	A algunos alumnos de mi clase les gusta ser los primeros.	3
24	Los alumnos queremos que nuestro trabajo sea mejor que el de nuestros compañeros.	3
33	En esta clase se favorece a algunos alumnos más que a otros.	3
51	Algunos alumnos forman pequeños grupos con sus íntimos amigos y no les importa el resto de los compañeros.	3
	FACTOR 4: Normas	
15	Los profesores nos enseñan a que respetemos las ideas y los sentimientos de las otras personas.	4; 8
16	Los profesores se muestran satisfechos cuando sacamos buenas notas.	4; 8
25	El aula es un lugar donde me siento solo.	4; 2

28	Los profesores se preocupan porque los alumnos no menospreciemos (o insultemos) a otros compañeros.	4
38	En esta clase los profesores esperan que los alumnos sigamos las normas.	4; 10
52	Creo que el tutor ha explicado claramente qué sucederá si un alumno rompe una norma.	4
53	Pienso que en esta clase existen demasiadas reglas y normas.	4
	FACTOR 5: Interés y preocupación	
21	Me parece que los profesores disfrutan con su trabajo.	5; 8
27	En esta clase los alumnos prestamos atención a lo que otros compañeros dicen.	5
29	Entre nosotros evitamos menospreciarnos (o insultarnos).	5
34	Las decisiones de esta clase son adoptadas por todos los alumnos.	5; 7
35	Los alumnos colaboramos muy bien entre nosotros.	5; 2; 7
39	En esta clase los profesores y los alumnos nos preocupamos unos de otros.	5
50	Los alumnos nos preocupamos mucho del progreso de esta clase.	5; 9
55	Creo que en esta clase se producen alborotos con frecuencia.	5
	FACTOR 6: Libertad	
17	Los profesores felicitan a los alumnos que ayudan a otros compañeros.	6; 8
36	En esta clase se propician debates.	6
37	En esta clase los profesores animan a los alumnos que quieren hacer las cosas de manera distinta.	6
54	Creo que en esta clase se producen alborotos con frecuencia.	6
	FACTOR 7: Cohesión	
34	Las decisiones de esta clase son adoptadas por todos los alumnos.	7; 5
35	Los alumnos colaboramos muy bien entre nosotros.	7; 2; 5
	FACTOR 8: Respeto y satisfacción	
14	Los profesores muestran respeto por nuestros sentimientos.	8; 1
15	Los profesores nos enseñan a que respetemos las ideas y los sentimientos de las otras personas.	8; 4
16	Los profesores se muestran satisfechos cuando sacamos buenas notas.	8; 4
17	Los profesores felicitan a los alumnos que ayudan a otros compañeros.	8; 6
21	Me parece que los profesores disfrutan con su trabajo.	8; 5
	FACTOR 9: Satisfacción e interés	
42	Los alumnos de esta clase nos interesamos por sacar buenas notas.	9
45	Creo que los profesores están satisfechos con «la marcha» general de los alumnos de este grupo de clase.	9
48	Los profesores se sienten orgullosos de esta clase.	9; 1
50	Los alumnos nos preocupamos mucho del progreso de esta clase.	9; 5
	FACTOR 10: Espacio y normas	
38	En esta clase los profesores esperan que los alumnos sigamos las normas.	10; 4
46	La clase es un espacio físico confortable.	10

Análisis estadístico

Como puede observarse, en la columna de la derecha de la Tabla I hay ítems que están en más de un factor. Considerando a priori que los factores resultantes no tenían por qué ser linealmente independientes u ortogonales, se optó por la rotación oblicua, dado que, además, la confluencia de estos ítems en más de una dimensión teórica del constructo Clima Social Aula era más acorde con la propuesta sustantiva.

En la Tabla II se aporta información sobre los tres distintos análisis que hemos realizado con la escala, todos ellos con el paquete estadístico SPSS 15, así como las medidas utilizadas, las decisiones previas y los indicadores para valorar las salidas resultantes. En los tres modelos se emplean como variables de entrada los ítems de la escala: se construye, para el caso del AF, la matriz de correlaciones de Pearson (no adecuada al nivel de medida ordinal), y en el EMD no métrico y en el AC las matrices de distancias euclídeas entre ítems.

TABLA II. Decisiones e indicadores de cada análisis

Análisis Factorial (AF)	Escalamiento Multidimensional (EMD) no métrico	Análisis Cluster (AC) jerárquico
Variables de entrada:		
Ítems 12-55 del cuestionario sobre percepción clima aula.	Ítems 12-55 del cuestionario sobre percepción clima aula.	Ítems 12-55 del cuestionario sobre percepción clima aula.
Matriz correlaciones Pearson entre ítems.	Distancias euclídeas.	Distancias euclídeas
Tipo de medida:		
No métrica (ordinal): escala Likert de cuatro puntos.	No métrica (ordinal): escala Likert de cuatro puntos.	No métrica (ordinal): escala Likert de cuatro puntos.
Correlaciones no adecuadas.	Distancias euclídeas.	Distancias euclídeas.
Decisiones previas:		
Componentes principales.	EMD no métrico (ALSCAL).	No introducir variables «inútiles».
Rotación oblicua.	Crear medida de disimilaridad (matriz de distancias euclídeas).	Crear medidas de disimilaridad: matriz de distancias euclídeas.

Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

Aumentar n.º iteraciones de convergencia mayor de 25.	Modelo ordinal de distancia euclídea con tres dimensiones.	Realizar método jerárquico para determinar dimensiones.
		Probar distintos métodos de conglomeración: se decide el método jerárquico de Ward.
Indicadores:		
Porcentaje de varianza.	*Stress* (entre 0 y 1).	Porcentaje de casos válidos/perdidos.
KMO.	R^2.	Dendrogramas combinando el método con la distancia «reescalada».
Esfericidad de Bartlett.	Gráfico de modelo de distancia euclídea.	
Determinante.	Gráfico de transformación de Shepard (valorar el ajuste).	
Matriz de configuración.		
Matriz estructura.		
Gráfico de sedimentación.		
Gráfico de componentes en espacio rotado.		

Entre las decisiones previas adoptadas (Tabla II), ha sido necesario aumentar en el AF el número máximo permitido de iteraciones de los algoritmos para encontrar la solución factorial. El programa recomienda por defecto no sobrepasar el valor de 25 para alcanzar un criterio de convergencia igual a cero. La solución que se obtiene inicialmente precisa de 43 iteraciones. Es necesario señalar que, en el caso que nos ocupa, se inicia el análisis con las respuestas de 407 alumnos, pero debido a las observaciones ausentes la muestra disminuye un 14,5%, eliminando hasta 59 casos por el procedimiento «lista» que aplica el programa por defecto (todas las variables afectadas por algún valor *missing*). En la Tabla III se muestran las opciones adoptadas sobre el número de iteraciones en combinación con diversos métodos de tratamiento de los valores *missing* y las matrices de entrada de correlaciones de Pearson y covarianzas, que son las que suelen emplearse habitualmente. Además de que ninguna de estas matrices es adecuada con medidas ordinales, tampoco la opción de excluir los valores ausentes según el procedimiento de lista permite el cumplimiento del criterio de convergencia. Lo más recomendable, por tanto, es tratar las no respuestas imputándoles el valor de la media de cada ítem, evitando así la supresión y disminución consiguiente de sujetos de la muestra, y cumpliendo, al mismo tiempo, el criterio de convergencia. Un estudio interesante para solventar la existencia de valores *missing* en análisis factoriales con datos no normales es el trabajo de Yuan, Marshall y Bentler (2002).

TABLA III. Número de iteraciones alcanzado según distintas opciones

		Tratamiento de datos *missing* (*)		
		(a) Lista	(b) Pareja	(c) Media
Matriz de entrada	Correlaciones	43	24	24
	Covarianzas	27	24	No converge en 50

En cuanto al EMD no métrico, se decide aplicar el procedimiento ALSCAL (*Alternating Least Squares SCALing*).

Respecto al Análisis Cluster, es sabido que los resultados se ven bastante afectados por la inclusión de variables «inútiles». Es recomendable, por tanto, atender a una cuidadosa selección de los ítems que se introducen siguiendo criterios sustantivos. Igualmente, se han probado distintos métodos de conglomeración y, por razones que se reflejarán en el siguiente apartado, el que mejor se ha ajustado ha sido el método de Ward.

Resultados

Análisis factorial

Como venimos diciendo, el AF de componentes principales requiere de condiciones (tales como normalidad, «homoscedasticidad» de las distribuciones de los ítems, presencia de relaciones lineales entre ellos), hipótesis previas que difícilmente se producen en la escala que empleamos. Sin embargo, algunos de los indicadores obtenidos, especialmente del AF, son francamente contradictorios ya que apuntan en la línea de que el Análisis Factorial se adapta a los datos (Tabla IV). Eso sucede, por ejemplo, con el determinante de la matriz de correlaciones que resulta próximo a cero, lo que lleva a suponer que existen ítems linealmente relacionados. Con la prueba de adecuación muestral KMO, por otro lado, se puede valorar si los datos se adecuan al modelo factorial: al presentar un valor de 0,906 cabe pensar que es una buena opción (en general se recomienda *KMO* > 0,80).

TABLA IV. KMO, prueba de Bartlett y determinante

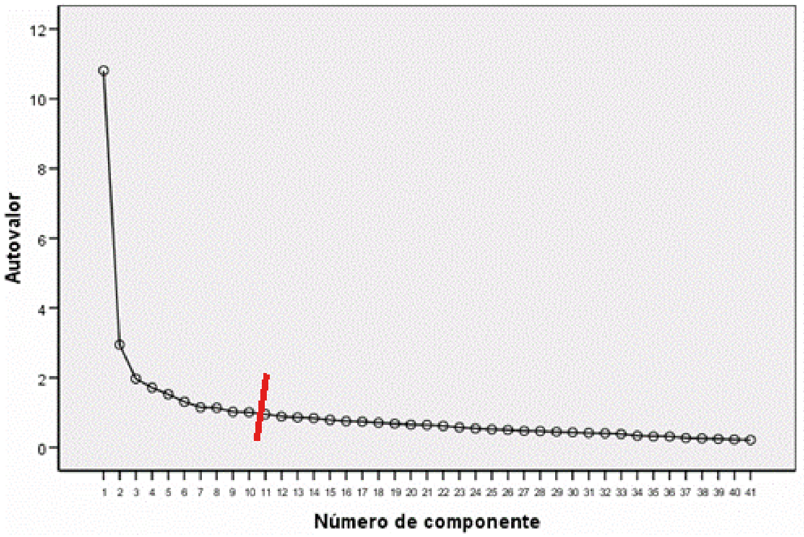
KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,906
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	5570,271
	gl	820
	Sig.	,000

Matriz de correlaciones[a]
a. Determinante = 5,30 E-008

En cuanto a la prueba de esfericidad de Bartlett, trabaja con el supuesto de que la matriz de correlaciones entre ítems es una matriz identidad, y en su contraste si $p < 0{,}05$, se rechaza la hipótesis nula de esfericidad, lo que supone que los ítems están correlacionados en la población de origen. Para ello se emplea una distribución de ji cuadrado ($\chi^2 = 5570$) que precisa de datos distribuidos normalmente, y con la escala ordinal no es posible. Luego, aunque se rechaza la hipótesis de esfericidad, y en contra de lo que parece de primera mano, no es un dato que ofrezca confianza. Además, la significación de la matriz de correlaciones que trabaja la prueba de Bartlett permite valorar sólo la presencia de relaciones no nulas, pero no el patrón de esas correlaciones (si están próximas a cero, si son o no son lineales, o su significación).

GRÁFICO I. Gráfico de sedimentación del AF

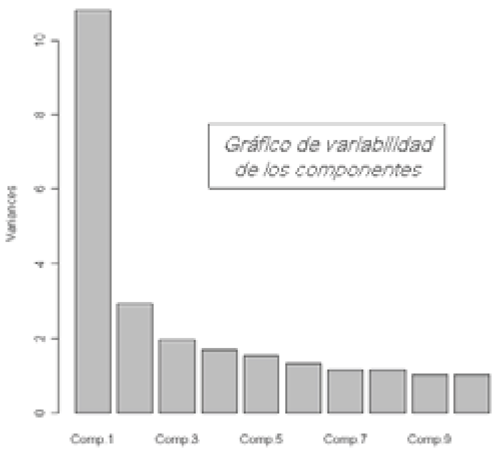


En el gráfico de sedimentación (Gráfico I) se representa el tamaño de los autovalores, lo que ayuda a determinar que el número óptimo de factores es diez, los correspondientes a autovalores mayores que 1 (criterio de Kaiser-Guttman). Este dato también puede señalarse con el punto de inflexión de la línea que dibuja el gráfico. Sucede, sin embargo, que la pendiente es muy pronunciada en los primeros componentes y muy suave en los últimos. Este efecto se aprecia bien en el Gráfico II: la desproporción de los porcentajes de varianza explicados es muy notable; sólo el primer factor explica más del triple del segundo (26,36 y 7,17, respectivamente –Tabla V–), y aunque el resto de los ocho componentes tienen autovalores mayores que 1, la explicación que aportan es mínima. Es evidente, por tanto, que en la escala que analizamos el gráfico de sedimentación y el tamaño de los autovalores no pueden considerarse criterios adecuados aunque cumplan con los estándares. Es claro que cuantas más variables hay, menos varianza necesita explicar un factor para alcanzar el criterio de Kaiser-Guttman (un factor con un autovalor de 1,0 explica el 10% de la varianza cuando hay 10 variables, pero sólo el 5% cuando hay 20 variables). Por esta razón, a pesar de su extendido uso, este criterio tiende a sugerir demasiados factores (Zwick y Velicer, 1986), y hay quienes no lo recomiendan (Cliff, 1988) o al menos sugieren que se

Revista de Educación, 354. Enero-Abril 2011, pp. 369-397
Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

acompañe con otros indicadores (Nunnally y Bernstein, 1995, p. 543) (Costello y Osborne, 2005).

GRÁFICO II. Variabilidad de los componentes en el AF



Por otro lado, el porcentaje de varianza total explicado asciende casi a 60% (Tabla V), pero la interpretación sustantiva de los diez factores que da la solución estadística es muy forzada. En la Tabla I hemos visto que los últimos factores quedaban configurados por un escaso número de ítems, siendo además multidimensionales (formando parte de otros factores previos) y aportando apenas explicación de la varianza total.

Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

TABLA V. Autovalores y porcentajes de varianza explicados en el AF

Varianza total explicada

Componente	Autovalores iniciales		
	Total	% de la varianza	% acumulado
1	10,808	26,361	26,361
2	2,943	7,178	33,539
3	1,972	4,809	38,348
4	1,714	4,180	42,528
5	1,521	3,709	46,237
6	1,309	3,193	49,430
7	1,146	2,795	52,225
8	1,134	2,767	54,992
9	1,019	2,485	57,476
10	1,006	2,454	59,930

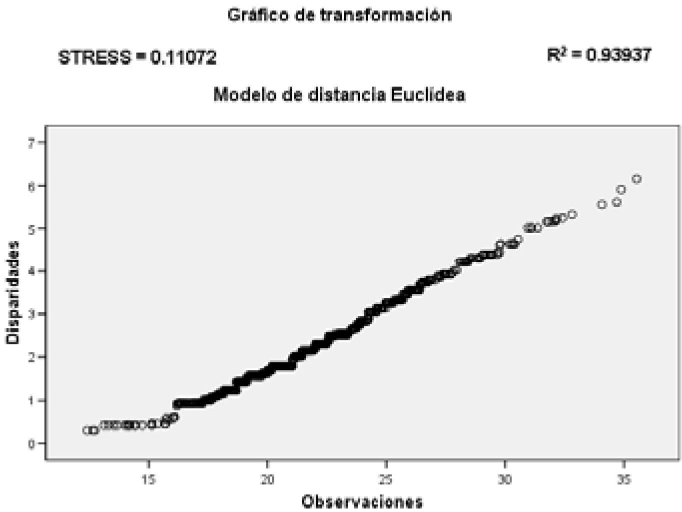
Métodos de extracción: Análisis de componentes principales.

Otro problema en los indicadores del AF se produce entre las matrices de configuración y de estructura. Al haber realizado una rotación oblicua ambas matrices no coinciden, con lo que debieran interpretarse tanto las saturaciones netas de la matriz de configuración como las correlaciones brutas de la matriz de estructura. No se incluyen aquí estas matrices por su amplia extensión, pero baste decir que se obtienen numerosos coeficientes entre ítems del mismo factor bajos o de signo contradictorio, y que la interpretación de los coeficientes respectivos no concuerda con la propuesta de los diez factores resultante de aplicar el criterio de Kaiser-Gutman.

Escalamiento Multidimensional no métrico

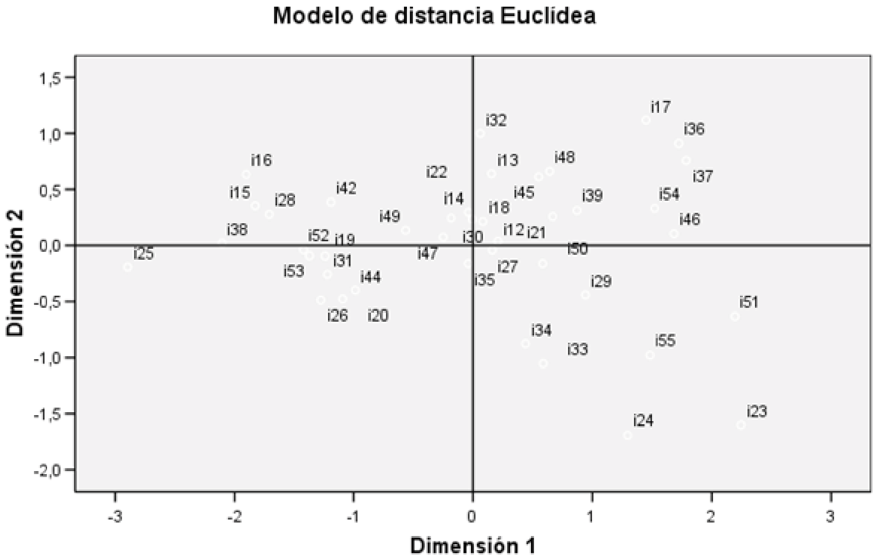
En el Gráfico III se muestran los indicadores obtenidos en el EMD no métrico. Valoramos, primeramente, la bondad de ajuste al modelo. El *stress* de Kruskal mide la diferencia entre las distancias euclídeas, que al tener un valor bajo –0,1107– está indicando un buen ajuste (Kruskal, 1964a y 1964b). La correlación múltiple cuadrada (proporción de varianza explicada por las distancias euclídeas entre los ítems) es alta –R^2= 0,9393–. El diagrama de Shepard representa los rangos originales de las proximidades (distancias euclídeas) frente a las distancias transformadas (disparidades) mostrando una función creciente suave y «sin saltos», reforzando la idea de un ajuste acertado.

GRÁFICO III. Ajuste de los datos al modelo de EMD no métrico y diagrama de Shepard



En el Gráfico IV se representan las distancias euclídeas estandarizadas en dos dimensiones. Si se toma la configuración de factores que resulta en el AF (Tabla I), pueden encontrarse ahora hasta ocho de los diez factores que se obtenían antes. La solución se ha reducido, es más parsimoniosa, y no se aleja demasiado de la solución sustantiva inicial –Tabla I–, excepto en los últimos factores que apenas explicaban variabilidad en el AF. Por tanto, si el EMD no métrico se adapta a la medida ordinal de la escala, los indicadores del ajuste son buenos y los mapas que resultan aclaran más la solución anterior del AF (igual sucede en los diagramas de tres dimensiones, aunque no estén aquí incluidos), es obvio que está aportando información complementaria de interés.

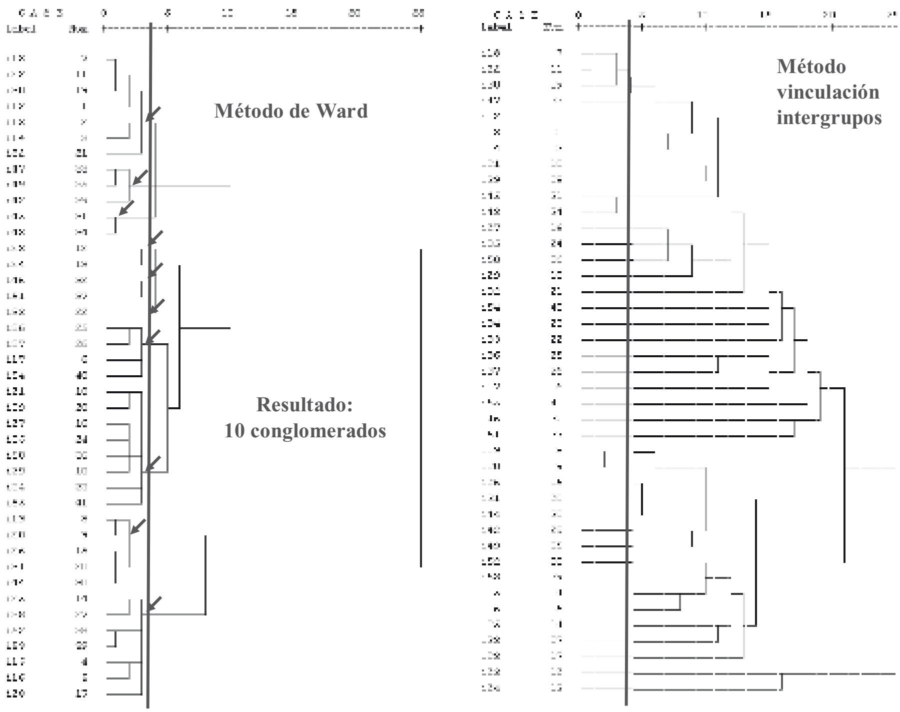
GRÁFICO IV. Mapa del EMD no métrico en dos dimensiones



Análisis Cluster

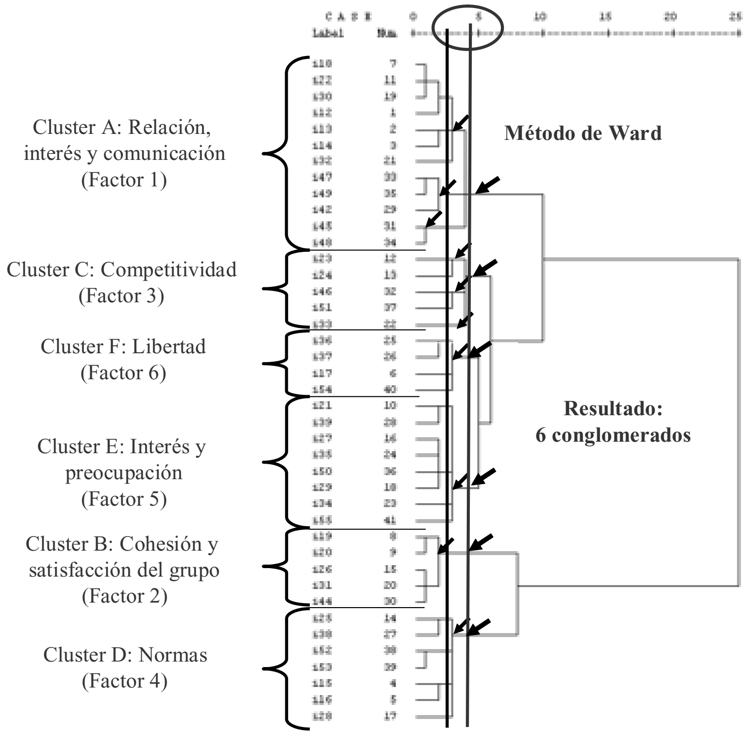
En el Gráfico V se muestran los dendrogramas de los Análisis Cluster realizados con los procedimientos de Ward y vinculación entre grupos. Es interesante observar cada una de las uniones que se construyen según una «distancia combinada o *'reescalada'*» mayor o menor. Las primeras agrupaciones se establecen con una distancia más pequeña, lo que indica que entre los ítems hay mayor similitud. Las respuestas a los ítems serán menos parecidas cuando sean incorporadas al conglomerado «más tarde», y en consecuencia con mayor distancia «reescalada». En ambos gráficos se ha trazado una línea vertical en el mismo valor de distancia. Puede observarse que la solución del método de vinculación intergrupos no es concluyente, mientras que con el método de Ward se consiguen diez conglomerados. Estas agrupaciones son, en alguna medida, similares a los factores obtenidos en el AF, pero en otra buena parte no, ya que con el AC no se permiten ítems multidimensionales, con lo que la solución de rotación oblicua del AF no es ahora compatible.

GRÁFICO V. Dendrogramas del AC jerárquico por los métodos de Ward y vinculación entre grupos



Ahora bien, lo realmente interesante es lo que ocurre en el Gráfico VI, donde hemos aumentado ligeramente la distancia combinada (moviendo la línea vertical hacia la derecha) mejorando la solución considerablemente: la pérdida por este pequeño incremento en la distancia «reescalada» se ve claramente compensada por la reducción del número de cluster de diez a seis, siendo estos grupos de ítems, además, más homogéneos internamente, por lo que el resultado aporta mayor confianza.

GRÁFICO VI. AC con el método de Ward y pequeño aumento de la distancia combinada



Visión integrada del conjunto de los resultados obtenidos

Recordando la configuración de factores del AF recogida en la Tabla I, y a la luz de los nuevos enfoques del EMD no métrico y del AC (Gráficos IV y VI), la nueva configuración de clusters e ítems cambia en algunos aspectos. Lo primero que se observa es la eliminación de los factores 7 a 10 del AF, formados mayormente por ítems multidimensionales que apenas explicaban varianza en el AF (Tabla V). El Cluster A (Relación, interés y comunicación) es el que muestra mayor número de ítems combinados con menor distancia (Gráfico VI), quedando conformado por los ítems que contenía el factor 1 del AF, pero incluyendo ahora los ítems 42 («Los alumnos de esta clase nos interesamos por sacar buenas notas»); 45 («Creo que los profesores están satisfechos con 'la marcha' general de los alumnos de este grupo de clase») y 49 («en este grupo aprendemos muchas cosas interesantes –a nivel personal y de conocimientos–»). En la salida del EMD no métrico estos ítems se encontraban muy próximos a la agrupación de los demás elementos de este grupo (Gráfico IV). En el Cluster B (Cohesión y satisfacción del grupo) se eliminan los ítems 25 («El aula es un lugar donde me siento solo») y 35 («Los alumnos colaboramos muy bien entre nosotros»), elementos que en la solución del EMD no métrico se situaban distantes al grupo de esta segunda dimensión (Gráfico IV). El ítem 25 vuelve a eliminarse nuevamente del factor en el que también estaba incluido (Cluster D –Normas–).

Discusión y conclusiones

Como hemos ido tratando, son numerosos los problemas detectados en el AF: (a) matriz de correlaciones inadecuada; (b) datos *missing*; (c) determinante de la matriz de rotación oblicua, prueba se esfericidad de Bartlett, prueba KMO y matrices de configuración y de estructura incoherentes; (d) incumplimiento del criterio de convergencia, etc. En una circunstancia habitual, y debido a que varios de estos indicadores cumplen con los criterios establecidos, podría haberse adoptado la solución factorial para interpretar la escala. Estaríamos, en ese caso, obligando a encajar los datos en el modelo estadístico, pero además eludiendo las características matemáticas del mismo.

Son varias las soluciones aportadas en las líneas anteriores, y algunas tan sencillas como matizar los cuantificadores lingüísticos de respuesta de la escala Likert; corregir los valores ausentes imputándoles la media del ítem; utilizar una matriz de correlaciones no paramétricas como matriz de entrada; desarrollar un Análisis Paralelo para determinar el número óptimo de factores o emplear otros procedimientos de estimación distintos a los componentes principales, como la estimación máximo verosímil. La solución que aquí hemos ejemplificado no ha hecho sino facilitar y mejorar la solución inicial factorial, y especialmente es relevante desde el punto en el que los dos modelos multivariantes empleados, el EMD no métrico y el AC jerárquico, se adaptan a la medida ordinal de la escala empleada.

Podemos señalar, no obstante, algunos matices. El EMD, por ejemplo, no aporta coeficientes de la configuración de los ítems, sino sólo coordenadas. Tampoco encontramos sustitución a las variables que genera el AF de componentes principales, puntuaciones que pueden emplearse en otros análisis posteriores: el AC jerárquico no proporciona valores de pertenencia a los clusters ni tampoco indicadores de la bondad del ajuste de los datos al AC.

Tanto unas razones como otras apuntan en la dirección de hacer un uso complementario de los tres modelos multivariantes de interdependencia cuando se tienen medidas ordinales: el EMD no métrico, el AC jerárquico y el Análisis de Componentes Principales no Lineal (ACPNL) (éste último, como mencionamos anteriormente, en sustitución al AF convencional, quedando pendiente su desarrollo explícito para un momento posterior). En todo caso, no se justifica el empleo exclusivo del AF de componentes principales lineal con una escala ordinal. Si el argumento que se esgrime es que los modelos multivariantes expuestos adolecen de menor potencia estadística, bien pueden empezar a explorarse las alternativas sustitutivas al AF también apuntadas: el análisis factorial de información completa implementado a partir de la TRI o el empleo conjunto del AF con un modelo de ecuación estructural. Son todos ellos modelos que respetan la métrica de los datos, y eso, desde nuestro punto de vista, es un argumento suficientemente importante.

Referencias bibliográficas

Arce, C. y Andrade, E. M. (2000). Recuperación de información métrica a partir de información no-métrica con diseños de escalamiento multidimensional incompletos. *Psicothema,* 12, 308-313.

Bartholomew, D. J. (2007). Three faces of factor analysis. En R. Cudeck & R. C. Maccallum (Eds.), *Factor analysis at 100. Historical development and future directions* (pp. 9-21). Mahwah, NJ: LEA.

Bartholomew, D. J., Steele, F., Moustaki, I. & Galbrain, J. I. (2002). *The analysis and interpretation of multivariate data for social scientists*. Boca Raton, FL.: Chapman & Hall/CRC.

Bernstein, I. H. (1988). *Applied Multivariate Analysis*. New York: Springer Verlag.

Bernstein, I. H. & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin,* 105, 467-477.

Bock, R. D., Gibbons, R. & Muraki, E. (1988). Full-information item factor analysis. *Applied Psychological Measurement,* 12, 261-280.

Bolton, R. J., Hand, D. J. & Webb, A. R. (2003). Projection techniques for nonlinear principal component analysis. *Statistics and Computing,* 13, 267-276.

Buja, A. & Eyuboglu, N. (1993). Remarks on parallel analysis. *Multivariate Behavioral Research,* 27, 509-540.

Cañadas, I. y Sánchez-Bruno, A. (1998). Categorías de respuesta en escalas tipo Likert. *Psicothema,* 10 (3), 623-631.

Cid, A. (2004). El clima escolar como factor de calidad en los centros de educación secundaria de la provincia de Ourense. *Revista de Investigación Educativa*, 22 (1), 113-144.

Cliff, N. R. (1988). The eigenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin,* 103, 276-279.

Cook, D. & Swayne, D. F. (2007). *Interactive and dynamic graphics for data analysis with R and Ggobi*. New York: Springer.

Costello, A. B. & Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assesment, Research & Evaluation,* 10, 1-9.

Fernández Díaz, M. J. y Asensio, I. (1993). Evaluación del clima de centros educativos. *Revista de Ciencias de la Educación,* 153, 69-83.

Ferrando, P. J. y Lorenzo, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test item: some aspects of the problem and some suggestions. *Psicológica,* 21, 301-323.

Flora, D. B. & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods,* 9, 466-491.

Franklin, S. B., Gibson, D. J., Robertson, P. A., Pohlmann, J. T. & Fralish, J. S. (1995). Parallel Analysis: A Method for Determining Significant Principal Components. *Journal of Vegetation Science*, 6, 99-106.

Gibbons, R. D., Bock, R. D., Hedeker, D., Weiss, D. J., Segawa, E., Bhaumik, D. K., Kupfer, D. J., Frank, E., Grochocinski, V. J. & Stover, A. (2007). Full-information item bifactor analysis of graded responsed data. *Applied Psychological Measurement,* 31 (1), 4-19.

Hair, J. F., Anderson, R. E., Tatham, P. L. & Black, W. C. (1999). *Análisis multivariante,* Madrid: Prentice Hall.

Jöreskog, K. G. (2007). Factor analysis and its extensions. En R. Cudeck & R. C. Maccallum (Eds.), *Factor analysis at 100. Historical development and future directions* (pp. 47-77). Mahwah, NJ: LEA.

Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, 1-27.

— (1964b). Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29, 115-129.

Little, R. J. A. & Rubin, D. B. (2002). *Statistical analysis with missing data.* New York: Wiley.

Muraki, E. & Carlson, J. E. (1995). Full-information factor analysis for polytomous item responses. *Applied Psychological Measurement*, 19, 73-90.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorial, and continuous latent variable indicators. *Psychometrika*, 49, 115-132.

Muthén, B. & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis and non-normal Likert variables: a note on de size of the models. *British Journal of the Mathematical and Statistical Psychology*, 45, 19-30.

Nunnally, J. & Bernstein, I. (1995). *Teoría psicométrica.* México: McGraw-Hill.

O'connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers*, 32, 396-402.

Pérez-Carbonell, A., Ramos, G. y López-González, E. (en prensa). Diseño y análisis de una escala para la valoración de la variable clima social aula en alumnos de educación primaria y secundaria. *Revista de Educación* (Madrid).

Picón, E., Varela, J. y Real, J. E. (2003). Clasificación y segmentación post hoc mediante el análisis de conglomerados. En J. P. Levy y J. Varela (Eds.), *Análisis multivariable para las Ciencias Sociales* (pp. 417-450). Madrid: Prentice Hall.

Real, J. E. y Varela, J. (2003). Escalamiento multidimensional. En J. P. Levy y J. Varela (Eds.), *Análisis multivariable para las Ciencias Sociales* (pp. 451-505). Madrid: Prentice Hall.

Rivas, T. y Martínez Arias, R. (1991). Relación entre escalamiento multidimensional métrico y análisis de componentes principales. *Psicothema*, 3, 443-451.

Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

Schriesheim, C. & Castro, S. (1996). Referent effects in the magnitude estimation scaling of frecuency expressions for response anchor sets: an empirical investigation. *Educational and Psychological Measurement*, 56, 557-569.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distances function (I y II). *Psychometrika*, 27, 125-139, 219-246.

Shigemasu, K. (1990). Factor analysis of categoriacal data. Recent developments. *Japanese Journal of Behaviormetrics*, 18, 41-51.

Swygert, K. A., Mcleod, L. D. & Thissen, D. (2001). Factor analysis for items or testlets in more than two categories. En D. Thissen & H. Wainer (Eds.), *Test Scoring* (pp. 217-250). Mahwah, NJ: LEA.

Thissen, D. & Wainer, H. (Eds.) (2001) *Test Scoring*. Mahwah, NJ: LEA.

Timm, N.H. (2002). *Applied Multivariate Analysis*. New York: Springer.

Yuan, K. H., Marshall, L. L. & Bentler, P. (2002). A unified approach to exploratory factor analysis with missing data, nonnormal data, and in the presence of outliers. *Psychometrika,* 67 (1) 95-121.

Zwick, W.R. & Velicer, W. F. (1986). Comparison of five rules for determining the Lumber of components to retain. *Psychological Bulletin, 99,* 432-442.

Dirección de contacto: Emelina López González**.** Universidad de Málaga. Facultad de Ciencias de Educación. Departamento de Métodos de Investigación e Innovación Educativa. Campus Teatinos S/N 29071 Málaga, España. E-mail: emelopez@uma.es

Fecha de entrada: 09-04-2008 Fecha de aceptación: 07-10-2008