Psicothema, 1995. Vol. 7, nº 3, pp. 655-665
ISSN 0214 - 9915 CODEN PSOTEG

SOFTWARE, INSTRUMENTACIÓN Y METODOLOGÍA

UN PAQUETE INTEGRADO EN SAS PARA COMPARAR SOLUCIONES FACTORIALES OBTENIDAS EN MUESTRAS DISTINTAS

Urbano Lorenzo Seva y Pere Joan Ferrando i Piera
Universitat Rovira i Virgili

Se presenta un paquete estadístico, escrito en el lenguaje matricial de SAS, que permite comparar patrones factoriales obtenidos al analizar las mismas variables en distintas muestras. El programa dispone de diversas alternativas que hacen posibles las comparaciones tanto entre patrones ortogonales como oblícuos. Se ha elaborado como un procedimiento interactivo de fácil utilización aún desconociendo el uso de SAS.

An IML SAS package for comparing factor patterns obtained in different samples. A statistical package written in IML SAS language is presented. The program allows the comparison between factorial patterns obtained when analysing the same variables in different samples. The implemented procedures allow the comparisons between both orthogonal and oblique patterns. It has been elaborated as an interactive procedure as easy to use that it is not necessary to know SAS package.

Cuando se utiliza el análisis factorial (AF) como una herramienta científica para asentar una teoría, posiblemente la condición central que debe alcanzarse es la de la invariancia en la solución obtenida. Uno de los aspectos de esta condición de invariancia, se refiere a que, si se analizan factorialmente las mismas variables en distintas muestras, entonces se espera la emergencia de los mismos factores. Thurstone expresa esta idea de forma muy clara:

Correspondencia: Pere Joan Ferrando i Piera
Departamento de Psicología.
Facultad de Psicología
Cra. de Valls, s/n. Apdo. Correos 567
43007 Tarragona. Spain

> " Si los factores primarios son los parámetros básicos y no meros reflejos de las condiciones experimentales o de las particulares condiciones de selección, entonces estos factores deben ser invariantes bajo condiciones selectivas que difieran ampliamente, y su interpretación debe ser la misma para los diferentes grupos experimentales." (Thurstone, L.L. 1950, p. 471)

Por supuesto la preocupación de Thurstone ha sido y sigue siendo compartida principalmente por investigadores sustantivos que pretenden generar modelos estructurales en base a los resultados obtenidos mediante AF. La invariancia es también un tema de capital importancia en estudios de tipo transcultural. En conjunto, este interés es el responsable de que los procedimientos para evaluarla constituyan un importante cuerpo de metodología dentro del AF.

Tras una revisión teórica bastante amplia, los autores consideramos que los métodos para evaluar la invariancia factorial se agrupan en tres categorías diferenciables: a) procedimientos para evaluar la similitud entre patrones factoriales; b) procedimientos para evaluar la similitud entre la puntuaciones factoriales estimadas y c) procedimientos de análisis de estructuras de covarianza simultáneamente en varios grupos. Sin entrar en polémica acerca de la superioridad de unos procedimientos sobre otros, diremos que el artículo que se presenta trata de métodos que pertenecen al primer grupo. Por otra parte cabe decir que la categorización es meramente ilustrativa y que no es rígida; en particular respecto a los grupos a) y c) algunos trabajos actuales (p. ej. McArdle y Cattell, 1994) parecen indicar una tendencia a la integración.

Justificación del trabajo

Aún cuando, de acuerdo con McDonald (1985), la necesidad de los métodos que aquí se presentan ha ido decreciendo al aparecer nuevos procedimientos, dado que el AF exploratorio sobre matrices de correlación es aún bastante popular entre investigadores aplicados pertenecientes a distintas áreas psicólogicas, los procedimientos que aquí se consideran pueden resultar de cierto interés. Desde nuestro punto de vista, estos métodos tienen su principal utilidad como etapas preliminares que ayuden a la formalización de hipótesis más elaboradas, las cuales permitirán utilizar modelos confirmatorios multigrupo.

Por otra parte, si bien es cierto que se han escrito rutinas informáticas para la mayor parte de los métodos aquí descritos, tales rutinas han aparecido citadas en revistas especializadas, se hallan muy dispersas y, siendo relativamente antiguas, pueden ser de muy difícil localización. Por estas razones creemos que su presentación en un programa integrado de fácil uso puede resultar conveniente.

Fundamentación teórica

Dentro de la invariancia de la solución factorial en el caso del análisis de las mismas variables en muestras distintas, se habla de una "invariancia configuracional" cuando las soluciones tienen la misma estructura aún cuando las saturaciones tengan valores numéricos distintos y de una "invarianza métrica" cuando los valores númericos de las saturaciones son los mismos en ambas muestras (Thurstone, 1950). Cattell denomina acertadamente "perfiles paralelos proporcionales" a aquéllos que cumplen la primera forma de invarianza (Cattell y Cattell, 1955). En el anexo I se ilustra este concepto.

Desde un punto de vista aplicado, debe considerarse que, debido a las fluctuaciones muestrales será casi imposible obtener un perfecto emparejamiento o una perfecta proporcionalidad entre los factores a comparar. Por tanto, los métodos que se presentan indicarán tan sólo hasta qué punto las soluciones son similares o proporcionales. Por otra parte la elección de una prueba de invarianza configuracional o métrica dependerá del hecho de que se mantengan o no las condiciones en que son medidas las variables (veáse Cattell y Cattell 1955 para una explicación detallada).

En esencia los métodos aquí presentados operan en la siguiente forma: se dispone de dos patrones factoriales a comparar. Estos patrones resultan del análisis de las matrices de correlación entre las mismas variables obtenidas en dos muestras distintas. En una primera etapa se lleva a cabo una transformación de los patrones que los lleva a una posición en la que la similtud entre ellos sea la máxima posible. En una segunda etapa se evalúa su similitud de acuerdo a las dos formas de invarianza descritas.

Para evaluar ambas formas de invarianza hemos elegido dos coeficientes bastante conocidos. El primero de ellos es el denominado "coeficiente de congruencia factorial" descrito inicialmente por Burt (1948). El segundo es un índice de discrepancia que se describe en Mulaik (1972). Las fórmulas para el cálculo e interpretación de estos índices se presentan en el Anexo I.

Respecto a los procedimientos para llevar a los patrones a comparar a la posición en que su similtud sea máxima, nos serviremos para su exposición del diagrama que se presenta seguidamente. Asimismo se dará tan sólo una descripción conceptual de los métodos. Su justificación detallada se presenta en el anexo II.

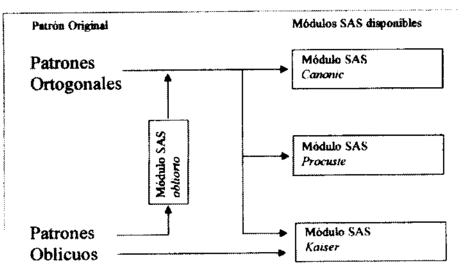

Figura 1. Métodos de comparación de soluciones factoriales disponibles en el procedimiento COMSOFAC (Ver texto)

Los procedimientos desarrollados en este trabajo se refieren al caso en que quieran compararse patrones factoriales obtenidos en dos muestras distintas. En general los métodos no exigen que el número de factores sea el mismo en ambas muestras, no obstante, en opinión de los autores es bastante conveniente que así sea. En principio, se dispone pues de dos patrones a comparar que pueden haber sido obtenidos tras extracción factorial directa (quizás mediante métodos distintos) o bién tras rotación ortogonal u oblícua.

Si los dos patrones a comparar corresponden a soluciones ortogonales (sea por extracción directa, sea por rotación rígida), entonces la metodología aquí seleccionada permite utilizar tres procedimientos distintos. El primero de ellos, descrito por Harman (1980) consiste en llevar a los dos patrones a la forma canónica mediante rotación ortogonal. Si B es el patrón rotado, el criterio de forma canónica implica que el producto B'B es una matriz diagonal. Las columnas de B pueden ordenarse de tal forma que los valores diagonales de la matriz producto se ordenen de mayor a menor. Una vez los patrones han sido rotados y reordenados, cada par de columnas de saturaciones equivalente en una y otra solución pueden ser comparados utilizando los dos índices anteriormente descritos.

Antes de pasar a la segunda alternativa, puede ser conveniente advertir que las soluciones directas de ejes principales (PAF) y de resíduo mínimo (MINRES) siempre están en forma canónica. No así las obtenidas tras extracción directa utilizando otros métodos.

La segunda alternativa está basada en el método de rotación Procustes ortogonal y se describe con detalle en diversos artículos (Cliff, 1966; Schönemann, 1966, Ten Berge, 1977; Brokken 1983). De las diversas variantes del método hemos optado por la más simple: una de las dos matrices a comparar se deja como patrón, mientras que la segunda se rota ortogonalmente a la posición de mejor ajuste en el sentido de mínimos cuadrados. Una vez obtenida la posición óptima, de nuevo los pares correspondientes de factores pueden compararse mediante los dos índices descritos.

La tercera alternativa es el método de Kaiser, Hunka y Bianchini, (KHB) descrito en un artículo de estos autores publicado en 1969 y, de nuevo, en 1971. Este procedimiento se lleva a cabo habitualmente sobre patrones correspondientes a soluciones oblícuas, pero puede emplearse también en el caso ortogonal.

En el caso de que uno o ambos patrones a comparar correspondan a soluciones oblícuas, la metodología utilizada permite dos alternativas. La primera de ellas consiste en devolver a los patrones a una forma ortonormal y, posteriormente utilizar cualquiera de las tres opciones antes descritas. La segunda alternativa consiste en utilizar el método de Kaiser Hunka y Bianchini directamente sobre los patrones oblícuos.

En el método KHB se considera un espacio común cuya base la forman los factores ortogonales correspondientes a uno de los dos estudios y en el que se situan las variables, factores ortogonales y factores oblícuos en ambos estudios. En el espacio definido se lleva a cabo la rotación de variables y factores de uno de los estudios (el que no se utiliza como base), bajo el criterio de que, cada par de vectores que representa a la misma variables en cada uno de los dos estudios tenga una distancia mínima; esto es: cada variable está representada por dos vectores, uno de ellos obtenido en una muestra y otro en la otra; se trata de que estos vectores estén lo más cercanos posible y que este criterio se cumpla para todos los pares de vectores. Lo que se pretende, por tanto, es que las mismas variables en uno y otro estudio coincidan lo mejor posible. Una vez determinada la transformación, se obtienen los cosenos de los ángulos entre factores, que se interpretan (aunque estrictamente no lo sean) como correlaciones. Lógicamente los cosenos de mayor interés son los que, supuestamente, corresponden a los mismos factores en cada una de las muestras. Si para un par de factores (supuestamente equivalentes) el coseno tiende a 1, entonces es razonable interpretar que este factor es el mismo en ambas muestras.

Para terminar esta exposición teórica, dado que este trabajo es principalmente de instrumentación no se entrará en aspectos de discusión acerca de los méritos o defectos de los distintos métodos disponibles. Para el lector interesado, diremos que pueden hallarse resultados comparativos entre el Procustes ortogonal y el método de KHB en Barrett (1986) y en Bijnen y Poortinga (1988). Asimismo en Reynolds y Harding (1983) se comparan los indicadores de similitud aquí utilizados con otros también de amplio uso.

Descripción del programa COMSOFAC.EXE

Los lenguajes de programación en los que se ha desarrollado el procedimiento de comparaciones de soluciones factoriales son el Turbo C y el lenguaje matricial IML integrado en el paquete estadístico SAS.

Mediante el primero lenguaje hemos desarrollado un programa (COMSOFAC.EXE) ejecutable del sistema operativo DOS. Este programa tiene como finalidad escribir el procedimiento SAS (COMSOFAC.SAS) que a continuación controla la comparación de las soluciones. Mediante el segundo lenguaje hemos desarrollado una librería de procedimientos que son los que realizan de hecho la comparación de las soluciones.

El procedimiento de comparación de soluciones factoriales que hemos desarrollado presenta las siguientes características:

a).- Es un procedimiento interactivo de fácil utilización incluso para usuarios no habituados a trabajar con el paquete estadístico SAS. El usuario no precisa conocer ni aprender la utilización de este paquete para realizar comparaciones. Por otra parte, es lo suficientemente flexible como para que el usuario experto pueda introducir las modificaciones que crea convenientes. Esta característica es esencial ya que el procedimiento no sólo se dirige a los expertos del AF, sino especialmente a los psicólogos aplicados que requieren herramientas eficaces, cuya utilización no conlleve unos conocimientos técnicos de software excesivos.

b).- Permite la comparación de soluciones que difieren tanto en número de factores como en número de variables.

c).- Las comparaciones pueden ser entre soluciones ortogonales y soluciones oblicuas.

d).- Las comparaciones de soluciones oblicuas pueden ser comparadas tanto en su posición oblicua como en su posición ortogonal previamente transformada.

El procedimiento en sí es una librería de módulos del lenguaje matricial IML. Estos módulos y su finalidad son las siguientes:

OBLIORT: transforma soluciones oblicuas a en soluciones ortogonales.

KAISER: realiza comparaciones siguiendo el método de Kaiser Hunka y Bianchini.

PROCUSTE: realiza comparaciones mediante el método procusteano.

CANONIC : realiza comparaciones mediante el método canónico.

Una representación gráfica de cómo se organizan estos módulos se puede observar en la figura 2.

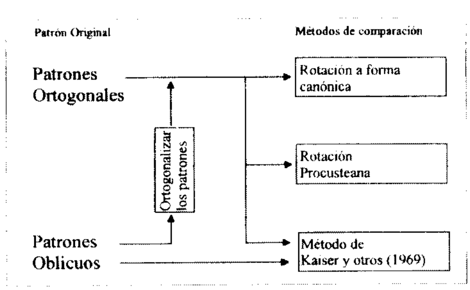

Figura 2. Módulos SAS disponibles en el procedimiento COMSOFAC (Ver texto)

Estos módulos son cargados y controlados por el procedimiento COMSOFAC.SAS. Este procedimiento presenta un menú en el que se puede configurar el tipo de comparación a realizar así como el método a aplicar.

No pretendemos describir como funciona cada módulo paso por paso, ya que por una parte resultaría repetitivo respecto a lo ya explicado en la fundamentación teórica del método y, por otra, caeríamos en explicar detalles excesivamente técnicos. Por el contrario, presentaremos cómo se utiliza el procedimiento y cómo se organiza el output de los estadísticos obtenidos.

Utilización del programa

El usuario conocedor del paquete estadístico SAS podría escribir por sí mismo el procedimiento COMSOFAC.SAS. Sin embargo, éste puede resultar un trabajo tedioso, ya que para cada par de soluciones a comparar los cambios a realizar en el procedimiento son numerosos. Por otra parte, para el usuario inexperto es un trabajo ciertamente complejo. Es por este motivo que hemos desarrollado el programa COMSOFAC.EXE. Este programa se controla por un menu interactivo que finaliza con la escritura del procedimiento COMSOFAC.SAS.

Previamente a la ejecución de COMSOFAC.EXE es preciso crear cuatro ficheros en código ASCII. En estos ficheros se ha de escribir la primera solución factorial (A1), la segunda solución factorial (A2), la matriz de transformación de la primera solución factorial (T1) y la matriz de transformación de la segunda solución factorial (T2), respectivamente. En este caso se entiende por matrices de transformación aquéllas que llevan desde una solución ortogonal a una solución oblícua.

Las matrices de A1, A2, T1 y T2 se pueden obtener directamente del output del paquete estadístico con el que se ha llevado a cabo el análisis. En el caso de soluciones ortogonales, la matriz de transformación será una matriz identidad. Si no se dispone de las matrices de transformación, pueden generarse dos ficheros ASCII alternativos con la matriz de correlaciones entre factores de la primera solución (PHI1) y la matriz de correlaciones entre factores de la segunda solución (PHI2). Cabe indicar que en este último caso el procedimiento KHB no podrá utilizarse.

Una vez se hayan creado estos ficheros se ha de ejecutar COMSOFAC.EXE. El menú interactivo que se obtiene se puede ver en la figura 3.

COMPARACION DE SOLUCIONES FACTORIALES	
SOLUCION PRIMERA	SOLUCION SEGUNDA
Solución Factorial:	Solución Factorial:
Número de Factores:	Número de Factores:
Número de Variables:	Número de Variables:
Matriz de Transform:	Matriz de Transform:
Matriz de Correlaci:	Matriz de Correlaci:

Figura 3. Menú de configuración del programa COMSOFAC.EXE. Para más detalles ver el texto.

La posición del cursor nos indica cual es la información que el programa requiere (p. ej. nº de factores). Los datos que el programa precisa son los nombres de los ficheros ASCII de las matrices A1, A2, T1, T2, PHI1 y PHI2, así como el número de factores (columnas) y variables (filas) de las soluciones factoriales A1 y A2. Los datos que el programa pide se introducen directamente sobre el menú, lo que facilita la compresión de los datos que el programa requiere en cada momento.

El programa finaliza preguntando si queremos ejecutar directamente el procedimiento COMSOFAC.SAS. En el caso de que lo queramos ejecutar directamente desde la línea de comandos del DOS, debemos escribir:

C:\SAS\SASEXE COMSOFAC.SAS

Como en el caso anterior, COMSOFAC.SAS presenta un menú que nos permite configurar la comparación de las soluciones. El menú que se obtiene se puede observar en la figura 4. Las posibilidades de este menú son las de decidir:

a.- si queremos transformar las soluciones oblicuas a sus posiciones ortogonales previamente a la comparación de las soluciones.

b.- si queremos aplicar el método de comparación de soluciones de Kaiser Hunka y Bianchini.

c.- si queremos aplicar el método de comparación procusteano.

d.- si queremos aplicar el método de comparación canónico.

```
Command ===>
        COMPARACION DE SOLUCIONES FACTORIALES
    ORTOGONALIZACION DE LAS MATRICES            [ ]
    METODO DE KAISER Y OTROS                    [ ]
    METODO PROCUSTEANO                          [ ]
    METODO CANONICO                             [ ]
```

Figura 4. Menú de configuración del programa COMSOFAC.SAS. Para más detalles ver el texto.

El menú correspondiente a COMSOFAC.SAS se puede ver en la figura 4. En este menú, para activar cualquiera de las optiones que hemos descrito, basta escribir una 'X' entre los corchetes que se hallan al lado de cada opción. Al pulsar la tecla INTRO el cursor avanza directamente a los siguientes corchetes. Para iniciar la comparación se debe escribir 'RUN' en la línea de comandos en la zona superior de la pantalla (la tecla INICIO nos lleva directamente a esta posición de la pantalla).

Descripción del output

Una vez ejecutado el paquete estadístico SAS, se obtiene el fichero COMSOFAC.LST en codigo ASCII. En este fichero podemos hallar el output producido durante la comparación de las soluciones factoriales.

En el caso de que se haya configurado el procedimiento para realizar todos los cálculos disponibles, el output que se obtiene se organiza del siguiente modo:

a).- Matrices A1 y A2 ortogonalizadas
b).- Método de Kaiser Hunka y Bianchini
  - matriz de correlación entre factores
c).- Método procusteano
  - solución A1
  - solución A2 transformada
  - matriz de índices de congruencia de Burt
  - matriz de índices de discrepancia de Mulaik
d).- Método canónico
  - solución A1 transformada
  - solución A2 transformada
  - matriz de índices de congruencia de Burt
  - matriz de índices de discrepancia de Mulaik

Ejemplos de utilización

A modo de ejemplo replicaremos, en primer lugar, la comparación de las soluciones factoriales que se lleva a cabo en el anexo original de Kaiser Hunka y Bianchini (1969). Por una parte se utiliza un análisis centroide en 4 factores seguido de una rotación oblícua obtenida mediante métodos gráficos. Por otra una solución ortogonal Varimax en dos factores. Los análisis proceden de estudios distintos, siendo las variables coincidentes seis medidas antropométricas. Se pretende comparar la similitud entre los dos primeros factores en la solución oblícua y los dos factores ortogonales.

El output del programa cuando se especifica el método KHB es el siguiente:

COMPARACION DE SOLUCIONES FACTORIALES

METODO DE KAISER, HUNKA & BIANCHINI

MATRIZ DE CORRELACION ENTRE FACTORES

	FAC1	FAC2
FAC1	0.6065183	0
FAC2	0	0.8573967

Los coeficientes de correlación son los cosenos entre los vectores FAC1 y FAC2 en ambos estudios. En el trabajo original de Kaiser Hunka y Bianchini (1969) los valores obtenidos son 0.61 y 0.86 respectivamente, que serían los de nuestro ejemplo redondeando a dos decimales.

En segundo lugar presentamos un ejemplo más complejo y, posiblemente más similar a los análisis que realiza un investigador en Psicología; los datos han sido cedidos amablemente por el Dr. Jordi Tous y proceden de su tesis doctoral.

Se trata de una escala de motivación laboral que consta de 46 reactivos en formato Likert de 7 puntos. Los primeros trabajos con esta escala tendían a indicar una solución en 4 factores. El contenido de los tres primeros era aquél que, "a priori" se esperaba de la teoría: El factor I se denominó "necesidades existenciales", el factor II "necesidades de crecimiento" y el factor III "necesidades de relación".

En una investigación posterior se administró la escala a dos grupos de sujetos claramente diferenciados en cuanto a su actividad profesional. Los sujetos del grupo I (N=426) tenían profesiones manuales; en tanto que los sujetos del grupo II (N=347) ocupaban cargos directivos en empresas. En cada uno de los grupos se llevó a cabo un análisis factorial de máxima verosimilitud proponiendo una solución en 4 factores. Se tenía interés en evaluar si la estructura del instrumento psicométrico era la misma en ambos grupos.

El programa COMSOFAC utilizó está vez como inputs los patrones obtenidos en ambos grupos tras solución directa. Se utilizaron dos métodos de comparación: el procustes ortogonal y la comparación a forma canónica. El programa devolvió el siguiente output:

COMPARACION DE SOLUCIONES FACTORIALES
METODO PROCUSTES ORTOGONAL

INDICES DE CONGRUENCIA FACTORIAL DE BURT

FAC1		FAC2	FAC3	FAC4
FAC1	0.9636032	0	0	0
FAC2	0	0.9895613	0	0
FAC3	0	0	0.9807288	0
FAC4	0	0	0	0.932873

INDICES DE DISCREPANCIA FACTORIAL DE MULAIK

FAC1		FAC2	FAC3	FAC4
FAC1	0.0713678	0	0	0
FAC2	0	0.0429013	0	0
FAC3	0	0	0.0661177	0
FAC4	0	0	0	0.0784372

METODO DE COMPARACION DE SOLUCIONES CANONICA

INDICES DE CONGRUENCIA FACTORIAL DE BURT

FAC		FAC2	FAC3	FAC4
FAC1	0.985522	0	0	0
FAC2	0	0.9658089	0	0
FAC3	0	0	0.9747488	0
FAC4	0	0	0	0.849326

INDICES DE DISCREPANCIA FACTORIAL DE MULAIK

FAC		FAC2	FAC	FAC4
FAC1	0.0686815	0	0	0
FAC2	0	0.0732853	0	0
FAC3	0	0	0.0462429	0
FAC4	0	0	0	0.3362388

En ambos casos los resultados son muy similares y la interpretación es la misma: los tres primeros factores en particular resultan muy semejantes en ambos grupos, tanto en perfil (índices de Burt) como en elevación (índices de discrepancia). El acuerdo entre ambos procedimientos y la claridad en la solución nos llevaría a interpretar que la estructura factorial del inventario es esencialmente la misma en ambos grupos.

Nota: El procedimiento COMSOFAC queda a disposición de los usuarios que los soliciten. Para ello deben dirigirse a los autores facilitando un disco y su dirección. Asimismo pueden solicitarse los listados de las rutinas IML.

Referencias

Barrett, P. (1986) Factor comparison: An examination of three methods. *Personality and individual differences*. 7, 327-340

Bijnen, E.J. y Poortinga, Y.H. (1988) The questionable value of cross - cultural comparisons with the Eysenck personality questionnaire. *Journal of cross-cultural Psychology*. 19, 2, 193-202

Brokken, F.B. (1983) Orthogonal procustes rotation maximizing congruence. *Psychometrika*. 48, 3, 343-349

Burt, C. (1948) The factorial study of temperamental traits. *British Journal of Psychology*. 1, 178-203.

Cattell, R.B. y Cattell, A.K.S. (1955) Factor rotation for proporcional profiles: Analytic solution and an example. *British Journal of Statistical Psychology*, 8, 83-92

Cliff, N. (1966) Orthogonal rotation to congruence. *Psychometrika*. 31, 33-42

Harman, H. (1980) *Análisis factorial moderno*. Madrid. Saltés

Kaiser, H.H.; Hunka, S. y Bianchini, J.C. (1969) Relating factors between studies based upon different individuals. En Eysenck y Eysenck (eds) *Personality structure and measurement*. London. Routledge

Kaiser, H.H.; Hunka, S. y Bianchini, J.C. (1971) Relating factors between studies based upon different individuals. *Multivariate Behavioral Research*. 6, 409-422

McArdle, J.J. y Cattell, R.B. (1994) Structural models of factorial Invariance in parallel proporcional profiles and oblique confactor problems. *Multivariate Behavioral Research*. 29, 1, 63-115

McDonald, R.P. (1985) *Factor analysis and related methods*. Hillsdale. LEA

Mulaik, S.A. (1972) *The foundations of factor analysis*. New York. McGraw-Hill

Reynolds, C.R. y Harding, R.E. (1983) Outcome in two large sample studies of factorial similarity under six methods of comparison. *Educational and Psychological Measurement*. 43, 723-728

Schönemann, P.H. A generalized solution to the orthogonal procustes problem. *Psychometrika*. 31, 1-10.

Ten Berge, J.M. (1977) Orthogonal procustes rotation for two or more matrices. *Psychometrika*. 42, 2, 267-276

Thurstone, L.L. (1950) *Multiple factor analysis*. Chicago. Univ. of Chicago Press.

Aceptado el 30 de noviembre de 1994

ANEXO I

Indices utilizados para evaluar la similitud entre patrones

Consideremos dos patrones factoriales a comparar obtenidos sobre las tres mismas variables:

	f a	fb
V_1	0.2	0.4
V_2	0.4	0.8
V_3	0.3	0.6

Gráficamente esta ordenación de datos permite las dos representaciones que se muestran en la figura 5.

Por una parte, la figura 5.a es la representación en perfil; dado que las columnas factoriales f a y f b están linealmente relacionadas (las cargas en f b son al doble que en f a), los perfiles similares en forma, pero difieren en elevación. Esta es la situación que Cattell denomina "perfiles paralelos proporcionales".

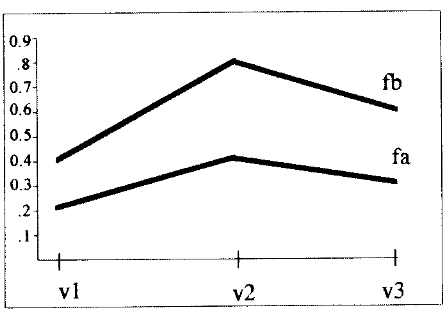

Figura 5.a. Representación de variables sobre factores (Ver texto)

Por otra parte, la figura 5.b representa las columnas fa y fb como vectores en el espacio de las tres variables. Nótese que, dada la proporcionalidad, los vectores son colineales, siendo fb más largo que f a.

El índice de congruencia factorial de Burt se obtiene por la expresión:

$$C_{ab} = \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2}} = \frac{0.2*0.4+0.4*0.8+0.3*0.6}{\sqrt{(0.2^2+0.4^2+0.3^2)*(0.4^2+0.8^2+0.6^2)}} = 1$$

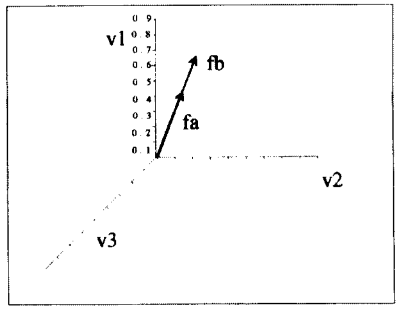

Figura 5.b. Representación de los factores sobre las variables (Ver texto)

Este índice es el coseno del ángulo entre fa y fb en la figura 5.b. Dado que son colineales, obviamente el coseno es 1. En la representación correspondiente a la figura 5.a, esto nos indica que ambos perfiles tienen el mismo patrón de elevaciones y descensos.

El índice de discrepancia viene dado por:

$$D_{ab} = \sqrt{\frac{\sum_{i=1}^{n}(a_i-b_i)^2}{n}} = \sqrt{\frac{(0.2-0.4)^2+(0.4-0.8)^2+(0.3-0.6)^2}{3}} = 31$$

En la figura 5.b es índice es la distancia entre los dos vectores. En la figura 5.a refleja la diferencia general en la elevación de los dos perfiles.

ANEXO II
Métodos utilizados en el procedimiento COMSOFAC

1.- Rotación a la forma canónica

Sea A una matriz patrón correspondiente a una solución ortogonal. Bajo el teorema fundamental del análisis factorial sabemos que,

AA' = Rc; (donde Rc es la matriz de correlación reproducida)

Se diagonaliza:

A'A = TDT' ; (Donde T es ortogonal y D diagonal)

Llevando a cabo ahora la rotación

B = AT

se obtiene

B'B=T' T D T' T = D; BB' = ATT'A; AA' = Rc

El patrón rotado B estará en forma canónica.

2.- Obtención de un patrón ortogonal desde un patrón oblicuo

Sea P un patrón oblicuo; de nuevo desde el teorema fundamental,

Rc = P F P'; Donde F es la matriz de correlación entre factores; dado que esta matriz es simétrica y de pleno rango es ortogonalmente diagonizable.

F=TD²T' (se escribe ahora D² por conveniencia)

P F P' = PTDDP'T'; haciendo B=PTD se tiene:
P F P'=BB'=Rc; B será un patrón ortogonal.

3.-Procustes ortogonal.

Sean A y B los patrones a comparar. Se busca una transfomación ortogonal de A.

A*=A*T, tal que tr((A*-B')*(A-B))= mínima.

Esta condición equivale a maximizar (véase Ten Berge, 1977):

tr(A*'B)=(T'A'B')= máxima.

La maximización de esta traza implica que la matriz producto T'A'B, sea simétrica y positiva semidefinida (SPSD)(véase de nuevo Ten Berge, 1977 para una demostración). Para alcanzar este requisito se lleva a cabo una descomposición de Eckart-Young que corresponde a:

A'B=PDQ' (Donde P y Q son ortogonales y D diagonal)

Se hace ahora

M=(A'B)(A'B)'=PDQ'QDP=PD²P'

M será simétrica de pleno rango, diagonalizando:

M=PLP' (L=D²), esta diagonalización permite obtener P

Ahora el producto:

P'(A'B)=P'PDQ'=DQ'; D⁻¹P'(A'B)=Q'

Finalmente haciendo T=PQ'.

El producto T'A'B = QP'PDQ'=QDQ' será una matriz SPSD; T será la matriz apropiada de transformación y A* = AT será el patrón ortogonal más similar a B en el sentido de los mínimos cuadrados.

4.K-H-B

Sean dos patrones rotados (ortogonales u oblicuos) P_1 y P_2 tales que

$P_1=A_1T_1$; $F_1=T'_1T_1$

$P_2=A_2T_2$; $F_2=T'_2T_2$

Donde A son los patrones procedentes de extracción factorial directa y T las matrices de transformación.
En primer lugar se normalizan los patrones originales:

$B_1=H_1$-1A_1; $B_2=H_2$-1A_2 (Donde $H_1$2 = diag (R_{1c})

Se pretende ahora llevar a cabo una rotación rígida:

$B^*_2= B_2T$; tal que B^*_2 sea lo más similar posible a B_1 en el sentido de los mínimos cuadrados.

La obtención de T se lleva a cabo exáctamente igual que en el apartado anterior. Obtenida T, el producto B_2T llevará a las variables del segundo estudio a la posición más cercana posible a las variables del primero en la misma forma, el producto T'_2T llevará a rotación a los factores del segundo estudio.
Por último, el producto T'_2TT_2 nos dará la matriz con los cosenos entre los factores del primer estudio y los factores del segundo tras la rotación que les lleva a la posición de máxima congruencia.