ALGUNAS RELACIONES ENTRE EL MODELO DE UN FACTOR COMUN Y EL MODELO LOGISTICO DE DOS PARAMETROS

Pere Joan FERRANDO PIERA y Urbano LORENZO SEVA
Universidad "Rovira i Virgili", Dpto. Educación y Psicología

En este trabajo se presentan las equivalencias entre las saturaciones factoriales en el análisis de un factor común sobre la matriz de correlaciones tetracóricas y los índices de discriminación y de dificultad en el modelo logístico de dos parámetros en la TRI. Dichas equivalencias se ilustran con un ejemplo numérico.
Palabras clave: Análisis factorial, Teoría de Respuesta a los Items, Análisis de Items.

Some relations between the one common factor model and the two parameter logistic model. The equivalence among factorial loadings in the one common factor model over the tetrachoric correlations matrix and discrimination and difficulty indexes in the two parameter logistic model are presented in this work. The relationships are illustrated with a numerical example.
Key words: Factor Analysis, Item Response Theory, Item Analysis.

De acuerdo con las teorías de Kuhn (1970), una de las características que debe poseer una disciplina para alcanzar su estado paradigmático es la de que exista un acuerdo generalizado en la terminología que utilizan los investigadores dedicados a su estudio. Es bien sabido que en Psicología aún no se ha alcanzado tal deseable estado de cosas y, dentro de ella, las áreas dedicadas al estudio de la medida no constituyen tampoco una excepción.

Los diversos modelos de medición en Psicología, suelen surgir como respuesta a un problema concreto (p. ej. la medida de la capacidad o de las actitudes) y desde marcos teóricos a veces bastante lejanos. En muchos casos, tanto los problemas como las soluciones son bastante similares y los modelos guardan estrechas relaciones entre ellos, sin embargo, al utilizar distinta terminología, la búsqueda de tales relaciones puede resultar bastante dificultosa.

En este trabajo trataremos de mostrar las relaciones existentes, tanto a nivel conceptual como matemático, entre dos modelos de notable importancia en la teoría de los tests: el modelo de un factor común de Spearman y el modelo logístico de dos parámetros perteneciente a la teoría de respuesta a los ítems. De acuerdo con una serie de autores (p. ej. Lord, 1980; Wainer y Messick, 1983; Takane y de Leeuw, 1987), se trata de dos modelos que surgen de una distinta tradición pero que se ocupan de unos tipos de datos muy similares y que guardan estrechas relaciones entre sí.

Por supuesto, existen una serie de publicaciones que se han ocupado ya del tema. Sin embargo, en muchos casos, estos trabajos se dedican exclusivamente a estudiar las relaciones entre los modelos de

Psicothema, 1993, vol. 5, nº 2, pp. 403-412
ISSN 0214 - 9915 CODEN PSOTEG

ojiva normal y los modelos de A.F. desarrollados por el grupo de Uppsala para el caso especial de variables dicotómicas. (Ver Christoffersson, 1975 y Muthén, 1982 para una revisión general de estos modelos). Adicionalmente, muchos de estos trabajos se ocupan de aspectos muy específicos, relacionados principalmente con métodos de estimación de parámetros. El presente trabajo, en cambio, pretende dar una visión más general, más simple y más aplicada, apoyándose en un conjunto de datos obtenidos en la adaptación de una escala de temperamento.

ORIGEN: BREVE REVISION HISTORICA Y SITUACION ACTUAL

Como es bien conocido, el modelo del análisis factorial (AF en adelante) se origina a principios de siglo con un famoso trabajo de Spearman (1904), en el que este autor proponía un método para verificar su hipótesis acerca de la estructura de la inteligencia. Desde el punto de vista psicométrico, el modelo estaba diseñado para trabajar con puntuaciones totales de tests y, para este tipo de datos, bastantes autores opinan que sigue siendo en la actualidad un método muy apropiado. (ver p. ej. Carroll, 1978; o McDonald, 1981). Este último autor, llega aún más lejos al afirmar que se trata del único modelo que los psicómetras entienden lo suficientemente bien como para garantizar su buen funcionamiento en situaciones prácticas (McDonald, 1982).

Cuando en la evolución de la teoría del test, la escala fue dejando paso al ítem como unidad de análisis, algunos investigadores trataron de incorporar la metodología del AF al análisis de ítems. En la década de los 40 fue quizás Lawley el principal autor que trató de relacionar y desarrollar simultáneamente ambos modelos de análisis.

Sin embargo, en aquella época la mayor parte de los reactivos eran de tipo dicotómico y, aparte de que para este tipo de variables el A.F. no puede ser legítimamente utilizado (ver McDonald 1985 para una explicación simple), su utilización parecía dar lugar a resultados distorsionados.

La inadecuación del modelo llevó a dos planteamientos distintos. Por una parte, algunos investigadores plantearon el concepto de los "factores de dificultad" y se preocuparon por la búsqueda del indicador de asociación apropiado en el caso de variables dicotómicas. Por otra parte, autores más radicales consideraron simplemente que el AF era un modelo inadecuado para este tipo de variables y que, en su lugar, debían plantearse modelos alternativos para el análisis de ítems binarios (Guttman, 1950).

Esta última postura fue defendida por autores cuya línea de investigación estaba vinculada a la medición de actitudes, los cuales propusieron los conocidos modelos del escalograma perfecto (Guttman, 1950) y del análisis de la estructura latente (Lazarsfeld, 1950). Aún cuando la actualmente denominada "Teoría de Respuesta a los Items" tiene también otras raíces, lo cierto es que estos últimos modelos constituyen una de sus bases más importantes.

Desde este resumen, quizás excesivamente simplista, se perfila por una parte, una tradición "factorialista", ligada principalmente a la Psicología Diferencial y unos modelos alternativos que surgen vinculados a otras áreas tales como la Psicología Social. Aún cuando las semejanzas entre estos modelos son mucho mayores que sus diferencias,(lo cual es actualmente reconocido por la práctica totalidad de los autores en ambas tradiciones), la distinta terminología utilizada propicia un cierto distanciamiento como hemos comentado al principio del artículo.

Antes de tratar de las relaciones entre el análisis de factor común aplicado a ítems binarios y el modelo logístico de dos parámetros puede ser necesario tratar de sintetizar las posturas actualmente adoptadas desde el análisis factorial.

a) En muchos casos se analiza directamente mediante AF la matriz de correlaciones producto-momento (coeficientes phi) considerando que, aunque el modelo no sea legítimo, las distorsiones que se producirán en el caso de que p se mueva en el rango 0.2 - 0.8 serán de poca importancia.

b) Es también frecuente considerar que las dicotomías constituyen una categorización artificial de un contínuo de respuesta subyacente. En este caso, el modelo AF sería aplicable legítimamente a dichas variables contínuas. Dado que la correlación tetracórica es el estimador de la correlación producto-momento entre variables artificialmente dicotomizadas, la práctica habitual consiste en llevar a cabo el análisis sobre la matriz de r-tetracóricas.

c) Bajo este mismo supuesto se desarrollan modelos de AF destinados al análisis de variables categóricas. En la misma forma que en el caso anterior, estos modelos pretenden obtener las saturaciones de las variables contínuas subyacentes a las dicotomías en los factores. Los principales representantes de este enfoque son Muthén y Christoffersson.

d) Otra perspectiva se basa en la consideración de que, mientras que el factor o variable latente constituye un contínuo ilimitado, la dicotomía tiene un rango de valores de frecuencia relativa necesariamente limitado entre cero y uno, por lo que, inevitablemente, en muchos casos las relaciones entre ambas serán no lineales. Desde esta postura se considera que la función logística constituye una buena aproximación a las regresión ítem-factor y se desarrollan modelos factoriales de tipo polinomial en los que se obtengan curvas similares a las logísticas. El principal representante de esta postura es McDonald (1967,1981, 1982).

Merece la pena destacar que, en la práctica, las estimaciones factoriales obtenidas desde las tres últimas perspectivas resultan ser todas ellas muy similares (McDonald, 1982; Takane y de Leeuw, 1987).

En el trabajo que aquí se expone es la segunda de estas posturas la que se toma en consideración al estudiar las equivalencias entre el modelo factorial y el modelo logístico.

RELACION ENTRE LOS MODELOS ESTRUCTURALES

El modelo estructural de un factor común, desarrollado inicialmente por Spearman (1904) puede escribirse en la forma:

$$xi = \bar{x} + l\theta i + e \tag{1}$$

Donde:

- xi: es la puntuación observada del individuo "i" en la variable "x"
- $\bar{x}$: es la media de la variable "x"
- l: es la saturación (coeficiente de regresión) de la variable "x" en el factor común
- θ_i: es la puntuación del sujeto "i" en el factor común o variable latente. El modelo supone que estas puntuaciones factoriales se distribuyen normalmente con media cero y varianza uno
- e: es el término de error, es decir, la parte de la puntuación observada que no puede explicarse por la influencia del factor.

Sobre los términos de error se realizan los habituales supuestos restrictivos que simplifican el modelo. Se asume que son independientes, tanto del factor como entre sí, y que se distribuyen con media cero.

Por último, el modelo considera que las distribuciones marginales, tanto de las variables observables como del factor, son normales. Adicionalmente se asume que las variables observables se distribuyen conjuntamente en forma normal multivariada.

Bajo estos supuestos, para un nivel fijo en el factor xi, las puntuaciones observadas en la variable x se espera tengan la siguiente media:

$$\bar{x} / \theta i = \bar{x} + 1\theta i \qquad (2)$$

Es decir, la media en la variable x condicionada al nivel xi en el factor, se espera igual a la media general más el producto de la saturación por el nivel fijo en el factor, ya que la media esperada en los errores es cero.

En el caso de que x sea un ítem dicotómico, entonces la media condicionada a un nivel fijo en el factor equivale, como es sabido, a la frecuencia relativa de aciertos en este ítem bajo este nivel. Esta frecuencia relativa puede interpretarse como la probabilidad de acertar el ítem, condicionada a un valor fijo en la variable latente, es decir:

$$P(x = 1 / \theta i) = \bar{x} + 1\theta i \qquad (3)$$

En el modelo logístico de dos parámetros, esta probabilidad condicionada viene dada por:

$$P(x = 1 / \theta i) = \frac{e^{Dax\,(\theta i - bx)}}{1 + e^{Dax\,(\theta i - bx)}} \qquad (4)$$

Donde:

D (1.7): es una constante utilizada para aproximar el modelo logístico al modelo de ojiva normal

bx: es el índice de dificultad del ítem. Se define como en nivel en el rasgo latente que corresponde al punto donde la función logística tiene pendiente máxima. Es fácilmente demostrable que, en este modelo, dicho punto corresponde a p=q = 0.5

ax: es el índice de discriminación del ítem. Puede definirse como la pendiente de la recta tangente a la función en el punto en que dicha pendiente es máxima (es decir p=q).

Para facilitar la equiparación entre (3) y (4), supondremos que en ambos modelos el factor o variable latente se distribuirá en forma normal con media cero y varianza uno. Es quizás este hecho el que permite una mayor distinción entre ellos, ya que en el modelo factorial, el factor necesariamente debe distribuirse normalmente en forma tipificada, mientras que en los modelos de la TRI no es imprescindible asumir la normalidad de la distribución de x (aunque puede ser conveniente) y, además, la métrica de x resulta en principio indeterminada. Esto nos llevaría a considerar que (4) es un modelo más general que (3) (Lord, 1980).

RELACIONES ENTRE LOS PARAMETROS DE AMBOS MODELOS

Para derivar las relaciones existente entre los parámetros de (3) y (4), seguiremos el desarrollo general que se lleva a cabo en Lord y Novick (1968, pp 377-378) y Lord (1980, pp 31-34). En estos trabajos se relacionan los parámetros del modelo de ojiva normal con los indicadores habituales en el análisis convencional de ítems desde la teoría clásica del test. Sin embargo, la línea general de razonamiento se adapta con facilidad a las relaciones que intentamos obtener aquí.

Consideremos en principio un ítem dicótomico "x". Supongamos que la puntuación en dicho ítem es una dicotomización artificial y que, bajo la respuesta puntuada, subyace una variable de respuesta "y" contínua y normalmente distribuida con media cero y varianza uno. Definamos un umbral "u" en la variable de respuesta tal que:

Si

$$y > u \Rightarrow x = 1$$

En otro caso

$$x = 0$$

Como hemos comentado antes, si "y" es contínua, el modelo factorial puede aplicarse legítimamente y, de acuerdo con (1).

$$yi = l\,\theta i + e \tag{5}$$

La representación gráfica de las relaciones hasta ahora expuestas se presenta en la figura 1

En este gráfico, "l", la saturación factorial, sería la pendiente de la recta de regresión de "y" sobre "x" , que equivaldría al coeficiente de correlación al estar ambas variables tipificadas.

Si la distribución condicional de las puntuaciones en "y" para un valor fijo de "x" es normal, tal como se representa en la figura, entonces el valor "zu" correspondiente a la ordenada que separa "p" (el área subrayada en la curva) de "q" , vendrá dado por:

$$Zu = \frac{u - y'}{s.\,e.} \tag{6}$$

Donde s. e. es la desviación de los valores de "y" en torno a la media condicional (error típico en terminología psicométrica). Como es bien conocido en teoría de la correlación, esta desviación, en el caso de variables tipificadas, se obtiene mediante:

$$s.\,e. = \sqrt{1 - l^2} \tag{7}$$

Como hemos visto en (2), "y'" sería la media condicional de "y" para el valor fijo "θi". Sustituyendo en (6) por la expresión equivalente:

Hagamos ahora:

$$ax = \frac{1}{\sqrt{1 - l^2}} \tag{8}$$

$$bx = \frac{u}{l} \tag{9}$$

Con un poco de álgebra es fácilmente demostrable que:

$$zu = ax\,(bx - \theta i) \tag{10}$$

Observando de nuevo la figura 1, puede observarse que el área sombreada "p" es decir, la probabilidad de acierto condicionada a θi vendrá dada por:

$$P\,(x = 1/\theta i) = \int_{ax(bx-\theta i)}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{\frac{-zu^2}{2}}\, dzu \tag{11}$$

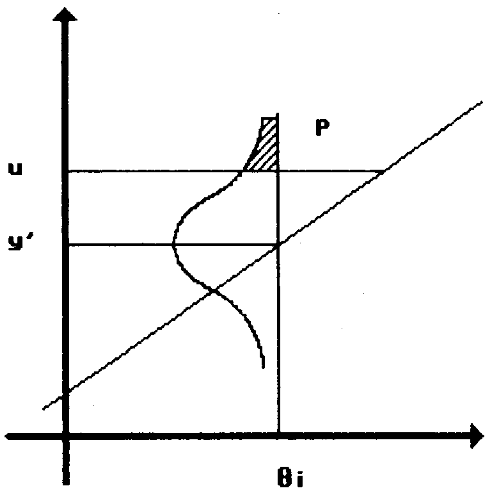

Figura 1. Relaciones supuestas entre la variable latente y la variable observable.

Y, puesto que la distribución es simétrica:

$$\int_{zu}^{\infty} = \int_{-\infty}^{-zu} \rightarrow \int_{ax\,(bx-\theta i)}^{\infty} = \int_{-\infty}^{ax\,(\theta i - bx)} \qquad (12)$$

De forma que (11) se convierte en la expresión del modelo de ojiva normal de dos parámetros cuyo equivalente logístico es (4). Como es conocido las diferencias en las estimaciones en ambos modelos difieren en menos de .001 (Birnbaum, 1968; Lord, 1980).

En principio pues, las saturaciones en el modelo factorial y los puntos de corte en la distribución de p y q en cada ítem (es decir, el valor z que separa la proporción de aciertos de la de errores bajo la curva normal) podrían ser útiles para predecir cuales serían los parámetros estimados por el modelo logístico. Sin embargo, estas equivalencias sólo pueden darse bajo una serie de condiciones restrictivas.

De entrada, las equivalencias mostradas relacionan el modelo factorial con el modelo de ojiva normal. Por tanto, las relaciones con el modelo logístico serán sólo válidas en cuanto la distribución logística constituya una buena aproximación a la normal (Takane y de Leeuw, 1987).

La exposición desarrollada en el apartado anterior nos indica la equivalencia entre la variable latente y un ítem particular. Sin embargo, tanto en el modelo factorial como en el logístico las estimaciones se llevan a cabo no para un ítem, sino para un vector de ítems (o, si se prefiere, para una matriz de sujetos x ítems). En ambos casos, para obtener las estimaciones deben hacerse una serie de supuestos restrictivos. Así, el método de máxima verosimilitud en análisis factorial estima las saturaciones bajo el supuesto de que los ítems se distribuyen todos ellos en forma normal multivariada y que la variable latente o factor se distribuye normalmente con media cero y varianza uno. Por otra parte, la estimación máximo verosímil marginal en el caso del modelo de ojiva normal, asume el mismo supuesto respecto a la variable latente y, de esta forma, la probabilidad conjunta de un vector de respuestas para todos los sujetos se obtiene integrando la probabilidad del vector de respuesta condicionada a un nivel de θ:

$$P(U/\theta) = \Pi\, p(\theta)^{ui}\, q(\theta)^{1-ui}\, g(\theta) \qquad (13)$$

(donde g(θ) es la función de densidad) a través de todos los valores de θ bajo la curva normal.

En suma, será sólo bajo los métodos de máxima verosimilitud en el análisis del factor común y de máxima verosimilitud marginal en el caso del modelo de ojiva normal, donde se encontraran las equivalencias hasta ahora descritas. La demostración formal puede hallarse en Takane y de Leeuw (1987). Cara a las pretensiones de este artículo, basta con las razones conceptuales: en ambos casos se utiliza el mismo criterio de estimación (máxima verosimilitud) y, también en ambos casos, los parámetros de los ítems se estiman condicionados al supuesto de que el rasgo latente se distribuye normalmente con media cero y varianza uno.

Por último, cabe observar de nuevo que las relaciones mostradas se refieren al análisis factorial sobre las variables contínuas que supuestamente subyacen a las dicotomías observadas y no al análisis factorial sobre las variables dicotómicas directamente. Dado que el análisis factorial usualmente se inicia no sobre la matriz de datos, sino sobre la matriz de correlaciones, las equivalencias se mantendrán en cuanto las correlaciones de la matriz a factorizar sean los estimadores máximo verosímiles de las correlaciones producto momento entre las supuestas variables subyacentes. Como es sabido dicho estimador máximo verosímil es la correlación tetracórica (Kendall y Stuart, 1982).

UN EJEMPLO NUMERICO

Para ilustrar las equivalencias expuestas en los apartados anteriores, hemos tomado los datos de una investigación realizada recientemente en nuestro laboratorio. En este trabajo se pretendía evaluar la adecuación del modelo logístico de dos parámetros a los reactivos de la subescala de impulsividad del EPI-A en su adaptación española (Seisdedos, 1973). Dicha subescala se compone de 10 ítems de los que se seleccionaron 6. El cuestionario fue administrado a una muestra de 2483 sujetos (1531 mujeres y 952 hombres). En relación al número de ítems, el tamaño muestral asegura la estabilidad de las estimaciones, tanto del modelo factorial como del modelo logístico.

En una primera etapa se obtuvieron los estadísticos convencionales en el análisis de reactivos. Cara a esta exposición nos interesan únicamente los índices de dificultad que se presentan en la tabla I

Items	2	3	4	5	6	10
p	.25	.19	.28	.32	.39	.48

Tabla 1. Indices de dificultad

En una segunda etapa se obtuvo la matriz de correlaciones tetracóricas inter-ítem mediante el programa PRELIS (Jöreskog y Sörbom, 1986). El algoritmo de dicho programa estima r-tetracórica mediante el método de máxima verosimilitud. Adicionalmente, se obtuvieron también las estimaciones de r-tetracórica mediante el procedimiento en series de potencias propuesto originalmente por Pearson, utilizando las 6 primeras potencias. Las estimaciones utilizando ambos métodos fueron muy similares para todos los valores de correlación.

En una tercera etapa, las matrices de correlación interítem fueron analizadas mediante el programa LISREL VI (Jöreskog y Sörbom, 1985) usando la aproximación del análisis factorial confirmatorio y proponiendo un modelo de un solo factor común. Como criterio de extracción se utilizó el de máxima verosimilitud. Las estimaciones obtenidas se presentan en la tabla II.

	Saturaciones	Unicidad
I2	.437	.809
I3	.850	.278
I4	.945	.107
I5	.295	.913
I16	.855	.270
I10	.418	.825

Tabla 2. Estimaciones del modelo factorial

En la cuarta etapa se parametrizaron los ítems de acuerdo al modelo logístico propuesto. Los parámetros fueron estimados mediante el programa MULTILOG (Thissen, 1986) que utiliza un algoritmo basado en el criterio de máxima verosimilitud marginal (Bock y Lieberman, 1970; Bock y Aitkin, 1981; Thissen y Steinberg, 1984). Respecto a este trabajo, el programa resulta apropiado en cuanto a que, en primer lugar la estimación es máximo verosímil y, en segundo lugar, los parámetros se estiman condicionados a que la variable latente se distribuya normalmente con media cero y varianza uno en el grupo normativo. Adicionalmente, recientes estudios de simulación con este programa parecen indicar que las estimaciones que proporciona son, generalmente, estables y precisas (Stone, 1992).

Los resultados de los ajustes se presentan en la tabla III.

En este punto, estamos en condiciones para verificar las equivalencias entre ambos modelos, de acuerdo con nuestra anterior exposición. Tomemos, por ejemplo, el ítem 2, cuya saturación estimada por el modelo factorial es de .437.

De acuerdo con (8), el índice de discriminación se obtendría por:

$$a = \frac{.437}{\sqrt{1 - .437^2}} = 0.486$$

Item			Valores estimados	
2				
I.	Discriminación (a) (S.e)	:	.478	(.069)
I.	Dificultad (b) (S. e)	:	1.530	(.144)
3				
I.	Discriminación (a) (S. e.)	:	1.601	(.170)
I.	Dificultad (b) (S. e)	:	1.049	(.040)
4				
I.	Discriminación (a) (S. e.)	:	3.662	(.952)
I.	Dificultad (b) (S. e)	:	.595	(.022)
5				
I.	Discriminación (a) (S.e)	:	.300	(.060)
I.	Dificultad (b) (S. e)	:	1.531	(.213)
6				
I.	Discriminación (a) (S. e.)	:	2.018	(.196)
I.	Dificultad (b) (S. e)	:	.33	(.025)
10				
I.	Discriminación (a) (S. e)	:	.453	(.063)
I.	Dificultad (b) (S. e)	:	.109	(.075)
	Chi-cuadrado		82.3	

Tabla 3. Estimaciones del modelo logístico

Mientras que en el ajuste de MULTILOG, el valor estimado es de 0.478

En la misma forma, el índice de dificultad de este ítem es de 0.25. El valor de z que separa 0.25 de 0.75 es de u=0.675.

El índice de dificultad vendría dado de acuerdo con (9) por: 0.675

$$b = \frac{0.675}{0.437} = 1.54$$

Siendo el valor estimado en MULTILOG de 1.53.

Los valores esperados desde el modelo factorial y las estimaciones del modelo logístico se presentan en la tabla IV.

Observando la última tabla, cabe observar que que los valores esperables desde el ajuste factorial se ajustan notablemente a las estimaciones del modelo logístico. Como indicador aproximado del ajuste, la

Item	/ par	Valores esperados	Valores estimados
2	a	0.48	0.47
2	b	1.54	1.53
3	a	1.61	1.60
3	b	1.03	1.05
4	a	2.90	3.66
4	b	0.61	0.59
5	a	0.30	0.30
5	b	1.59	1.53
6	a	1.65	2.01
6	b	0.33	0.33
10	a	0.45	0.45
10	b	0.10	0.11

Tabla 4. Valores esperados y valores estimados

correlación entre columnas es del orden de .9. Sin embargo, no consideramos necesario utilizar pruebas para comparar la similitud de las soluciones.

Es de destacar asimismo, que las distorsiones mayores se producen en los índices de discriminación de los ítems 4 y 6. En ambos casos, estos son los valores de pendiente más elevados y, asimismo, las estimaciones de MULTILOG que muestran un error típico más alto, lo cual no resulta en ningún caso sorprendente. Podríamos concluir diciendo que, en caso de ajustes con poca variabilidad de error, las relaciones esperadas se cumplen con bastante aproximación.

DISCUSION

El estudio de las relaciones de equivalencia entre modelos aparentemente distintos, (al menos desde el habitual punto de vista del investigador aplicado), puede resultar de cierta utilidad general en cuanto a que facilita una perspectiva más unificadora y justifica la utilización conjunta de distintas metodologías.

Más específicamente, en los estudios psicométricos en que se ajustan escalas de acuerdo a los modelos de TRI, es habitual evaluar previamente la dimensionalidad de los reactivos mediante análisis factorial, antes de proceder al ajuste del modelo. Esto es así en cuanto a que la unidimensionalidad es un requisito básico que deben cumplir los ítems que componen una escala.

Este trabajo puede proporcionar quizás alguna indicación acerca de cual puede ser el análisis de dimensionalidad apropiado que se realice previamente a la parametrización logística. Por razones que escapan a la finalidad del artículo, creemos que el análisis en componentes principales y las habituales pruebas asociadas (proporción de varianza asociada) resultan totalmente inapropiados en el caso del análisis de reactivos, lo que lleva necesariamente al uso de un modelo de factores comunes.

La elección del modelo comentado en este artículo, sin embargo, tiene también limitaciones. En primer lugar, si los niveles de dificultad de los reactivos son intermedios, es dudoso que la complejidad de la estimación de las tetracóricas justifique esta alternativa frente al análisis de la matriz de coeficientes phi obtenida directamente en cualquier programa estadístico.

Por otra parte, la utilización de estas correlaciones presenta algunos problemas sobradamente conocidos. Dado que la correlación tetracórica es un estimador de una relación normal bivariada (no de una distribución conjunta normal multivariada), en algunos casos la matriz obtenida será no gramiana, lo que exigirá una transformación que la convierta en semidefinida positiva y posibilite su factorización. Por otra parte, en los casos de valores extremos en la tabla de contingencias, los valores obtenidos pueden resultar bastante distorsionados o bien la convergencia de la estimación puede fallar.

En definitiva, cabe considerar que, aún cuando actualmente, la factorización de la matriz de correlaciones tetracóricas sea una de las alternativas más razonables para estimar la dimensionalidad de los ítems antes de proceder a su parametrización logística, las limitaciones propias del método hacen deseable el avance de los otros modelos alternativos que hemos comentado anteriormente.

REFERENCIAS

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In Lord, F.M. y Novick, M.R. *Statistical theories of mental tests scores.* Massachusetts. Addison-Wesley.

Bock, R.D. y Lieberman, M. (1970). Fitting a response model for n dichotomously scored items. *Psychometrika.* 35, 179-197.

Bock, R.D. y Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: an application of the EM algorithm. *Psychometrika.* 46, 443-459.

Carroll, J.B. (1978). How shall we study individual differences in cognitive abilities? *Intelligence.* 2, 87-115

Christoffersson, A. (1975). Factor analysis of dichomized variables. *Psychometrika.* 40, 1, 5-31.

Guttman, L. (1950). The problem of attitude and opinion measurement. En Stouffer, S.A. (ed) *Measurement and prediction* Vol IV. Princeton. Princeton Univ. press

Jöreskog, K.G. y Sorbom, D. (1985). LISREL VI: *Analysis of linear structural relationships by the method of maximum likelihood.* Moreesville, Ind. Scientific Software Inc.

Jöreskog, K.G. y Sorbom, D. (1986). PRELIS: *A program for multivariate data screening and data summarization: a preprocessor for* LISREL.Moreesville, Ind. Scientific Software Inc.

Kendall, M.G. y Stuart, A. (1982). *The advanced theory of statistics.* Vol II. London. Charles Griffin.

Kuhn, T.S. (1970). *The structure of scientific revolutions.* Chicago. Univ. Chicago Press.

Lazarsfeld, P.F. (1950). The logical and mathematical foundation of latent structure analysis. En Stouffer, S.A. (ed) *Measurement and prediction* Vol IV. Princeton. Princeton Univ. press

Lord, F.M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale. LEA

Lord, F.M. y Novick, M.R. (1968). *Statistical theories of mental tests scores.* Massachusetts. Addison-Wesley.

McDonald, R.P. (1967). *Non linear factor analysis.* Psychometric monograph n. 15

McDonald, R.P. (1981). Dimensionality of tests and items. *Journal of Mathematical and Statistical Psychology.* 34, 100-107.

McDonald, R.P. (1982). Linear vs. non linear models in Item Response Theory. *Applied Psychological measurement.* 6, 4, 379-396.

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

Muthén, B. (1982). Some categorical response models with continuous latent variables. En Jöreskog, K.G. (ed) *Systems under indirect observation* part I. Amsterdam. North Holland.

Seisdedos, N. (1973). *Cuestionario de personalidad E.P.I.* Madrid. TEA Ediciones.

Spearman, Ch. (1904). General intelligence; objectively determined and measured. American *Journal of Psychology.* 115, 201-292.

Stone, C.A. (1992). Recovery of marginal maximum likelihood estimates in the two parameter logistic response model: An evaluation of MULTILOG. *Applied Psychological Measurement.* 16, 1, 1-16.

Takane, Y. y de Leeuw, J. (1987). On the relationship between Item Response Theory and factor analysis of discretized variables. *Psychometrika.* 52, 3, 393-408.

Thissen, D. (1986). MULTILOG user's manual. Mooresville. Scientific Software

Thissen, D. y Steinberg, L. (1984). A model for multiple choice items. *Psychometrika.* 49, 501-519.

Wainer, H. y Messick, S. (1983). *Principles of modern psychological measurement.* Hillsdale. LEA