

SOFTWARE, INSTRUMENTACION Y METODOLOGIA

UTILIZACION DE MODELOS UNIDIMENSIONALES DE TEORIA DE RESPUESTA A LOS ITEMS CON DATOS MULTIFACTORIALES

Marcelino Cuesta y José Muñiz

Facultad de Psicología. Universidad de Oviedo.

Se presenta un trabajo con datos simulados en el que se trata de investigar el comportamiento de las estimaciones realizadas sobre datos multifactoriales a partir de los modelos unidimensionales de Teoría de Respuesta a los Items de uno y tres parámetros. La simulación se llevó a cabo siguiendo el modelo factorial de Wherry, Naylor, Wherry y Fallis (1965). Se generaron dieciséis bases de datos, variando el número de items, el número de factores y sus características. Los resultados indican que cuando existe un factor dominante ambos modelos tienden a estimar dicho factor. Si existen varios factores independientes el modelo de tres parámetros capta sólo uno de ellos, mientras que el modelo de un parámetro capta una combinación de todos los factores.

Palabras clave: Modelos logísticos; Unidimensionalidad; Robustez.

Use of unidimensional item response theory models with multifactorial data. A simulation study is presented in order to investigate the relationship between the factorial complexity and the robustness of parameter estimates from unidimensional Item Response Theory models of one and three parameters. The simulation was carried out in accordance with the Wherry, Naylor, Wherry and Fallis (1965) factorial model. Sixteen data sets were generated varying the number of items, number of factors and its characteristics. The results show that when a dominant factor is present this factor is collected by both models. When several independent factors are present, the three parameter model picks up only one of them, whereas the one parameter model collect a combination of all these factors.

Key words: Logistic models; Unidimensionality; Robustness.

Con el predominio teórico en Psicometría de los modelos de Teoría de Respuesta a los Items (TRI) a partir de los años 70, aparecen algunos interrogantes acerca de su adecuación a la práctica cotidiana. Una de las más importantes cuestiones surgidas en este ámbito ha sido la referida al cumplimiento del supuesto de unidimensionalidad, requerido por los más populares modelos de TRI. Es decir, la mayoría de estos modelos exigen que los datos sean unidimensionales, mien-

tras que como es bien sabido, los datos reales sobre los que se trabaja normalmente en la práctica no son nunca estrictamente unidimensionales. Esta preocupación dio lugar a una línea de investigación centrada en la robustez de los modelos unidimensionales de TRI a la violación de dicho supuesto. Como trabajo pionero en este campo puede considerarse el presentado por Reckase (1979), al que han seguido otros muchos (Ackerman, 1989; Ansley y Forsyth, 1985; Doody-Bogan y Yen, 1983; Drasgow y Parsons, 1983; Harrison, 1986; Way, Ansley y Forsyth, 1988; Yen, 1984).

En el presente trabajo se llevó a cabo una replicación parcial del estudio de Reckase

Correspondencia: Marcelino Cuesta
Facultad de Psicología. Universidad de Oviedo.
Aniceto Sela, s/n
33005 Oviedo - Spain

(1979). El objetivo perseguido era estudiar la influencia y posible distorsión que sobre las estimaciones de la habilidad de los sujetos, realizadas mediante modelos logísticos unidimensionales de uno y tres parámetros, tiene la complejidad factorial de los datos analizados. Para ello la lógica general consistió en simular datos con distinto grado de dimensionalidad y ver en qué medida afectaba esta dimensionalidad a las estimaciones de las puntuaciones de los sujetos hechas mediante los modelos logísticos unidimensionales de uno y tres parámetros.

METODO

Modelo de Simulación

Para la generación de los datos de acuerdo a estructuras factoriales previamente determinadas se empleó el método descrito por Wherry et al. (1965). Una vez decidido el número de factores a generar, el modelo básico de generación de los datos consiste en una matriz con los valores de los pesos factoriales deseados y un vector con los valores elegidos para el parámetro de dificultad clásico de cada ítem. Partiendo de la matriz de pesos factoriales el procedimiento genera una puntuación típica para cada sujeto en cada ítem. Esta puntuación se dicotomiza tomando como punto de corte la puntuación típica que deja por encima de sí la proporción de casos que indica el índice de dificultad asignado a ese ítem. De esta forma para el total de sujetos se obtiene una proporción de respuestas correctas aproximada a la especificada por el índice de dificultad establecido para ese ítem. La realización concreta de este proceso de simulación se llevó a cabo enteramente empleando el paquete estadístico SPSS/PC+, siguiendo las directrices marcadas por Bernstein (1988).

Bases de Datos

En la construcción de las diferentes bases de datos empleadas se manipularon tres as-

pectos: número de ítems, número de factores y características de los factores. Con respecto a la longitud de los tests simulados se utilizaron dos longitudes: $n=30$ y $n=50$. En cuanto al número de factores, para cada longitud de test se crearon bases que iban desde un único factor a un máximo de cinco factores. Para aquellos casos en los que la base debía contar con tres, cuatro o cinco factores se creaban dos tipos diferentes de estructuras: una en la que todos los factores tienen el mismo peso y otra en la que hay un factor claramente dominante.

Conforme a las especificaciones que se acaban de relacionar se generaron ocho bases para cada una de las dos longitudes de test manejadas. En todos los casos el tamaño de muestra fue de mil sujetos. Así, las dieciséis bases de datos quedaron configuradas con las características que se muestran en la tabla 1.

Tabla 1

Descripción de las bases de datos

Número de factores resultantes	Abreviación
Un Factor	$n=30$ (30-1F) $n=50$ (50-1F)
Dos Factores	$n=30$ (30-2F) $n=50$ (50-2F)
Tres Factores con uno Dominante	$n=30$ (30-3FD) $n=50$ (50-3FD)
Tres Factores Equivalentes	$n=30$ (30-3FE) $n=50$ (50-3FE)
Cuatro Factores con uno Dominante	$n=30$ (30-4FD) $n=50$ (50-4FD)
Cuatro Factores Equivalentes	$n=30$ (30-4FE) $n=50$ (50-4FE)
Cinco Factores con uno Dominante	$n=30$ (30-5FD) $n=50$ (50-5FD)
Cinco Factores Equivalentes	$n=30$ (30-5FE) $n=50$ (50-5FE)

Análisis

Para realizar los análisis necesarios de las 16 bases de datos se utilizaron el paquete estadístico SPSS/PC+ (Versión 3.0) y el paquete específico de TRI, PC-BILOG (Versión 1.1).

Como paso previo a la configuración definitiva de las bases arriba descritas, formadas por unos y ceros, se llevaron a cabo análisis factoriales de las variables continuas generadas por el modelo con el fin de comprobar que los resultados se ajustaban a lo diseñado por nosotros. Una vez efectuada la comprobación los datos fueron dicotomizados de la forma detallada con anterioridad.

Los análisis llevados a cabo sobre las bases de datos finalmente generadas pueden agruparse en tres bloques:

- Análisis factorial. Cada una de las bases fue sometida a un análisis de factores principales y a un análisis de componentes principales sobre correlaciones phi. En cada uno de ellos se calcularon las puntuaciones factoriales de los sujetos en los k factores que subyacen a las bases no dicotomizadas.

- Estimaciones de los parámetros de TRI. Empleando el programa BILOG se estimaron los parámetros de los sujetos para los modelos logísticos de uno y tres parámetros.

Se eligió el modelo de un parámetro por su amplia utilización por una parte de los usuarios de la TRI. Si bien es cierto que desde el punto de vista teórico ha sido bastante criticado por añadir dos supuestos bastante restrictivos (iguales discriminaciones y ausencia de azar) a los generales de la TRI, no es menos cierto que a nivel práctico ha sido el preferido por los usuarios por poseer unas características que lo hacen muy atractivo (Hambleton & Swaminathan, 1985; Muñiz, 1990):

- Incluye pocos parámetros y es fácil de trabajar con él (su lógica es muy simple, la respuesta de un sujeto depende únicamente de su nivel en la variable que estemos midiendo y de la dificultad del ítem).

- Los problemas de estimación de los parámetros son menores que en los otros modelos.

Añadir a lo dicho que parece, que aun cuando se violen algunas de las asunciones del modelo, éste ajusta razonablemente a los datos (Muñiz, Rogers y Swaminathan, 1989).

El otro modelo empleado, el de tres parámetros, se eligió por ser el más general de los modelos logísticos de más amplio uso.

- Análisis clásico. Se calcularon también los estadísticos descriptivos de cada muestra y una estimación de la fiabilidad por medio del coeficiente alfa.

Con los resultados obtenidos de los análisis que se acaban de presentar se creó, para cada base, una nueva matriz que fue nuevamente sometida a análisis correlacionales. Esta matriz estaba formada por las puntuaciones factoriales de los sujetos (PCK, PAFk), las puntuaciones empíricas (P.Emp.) y las puntuaciones θ estimadas con el modelo de uno y tres parámetros (θ_1, θ_3).

RESULTADOS

Antes de pasar a la presentación de los resultados obtenidos del análisis de la matriz descrita en el punto anterior, y que constituyen el núcleo central del trabajo, se presenta un resumen descriptivo de las bases de datos originales (tablas 2 y 3). Tanto en las muestras de 30 ítems como en las de 50 ítems para un mismo número de factores el primer autovalor toma valores notablemente más altos para las bases con un factor dominante, como era de esperar. También puede verse como no hay apenas diferencias entre los resultados obtenidos desde los análisis de componentes principales y los de factores principales. Un dato importante a señalar en este contexto son los altos valores que toma el coeficiente alfa. Si bien es cierto que baja muy levemente su valor cuando se trabaja con factores independientes, aun en esos casos alcanza niveles que en el peor de los casos se pueden calificar como altos. Este

Tabla 2
Descripción de las bases de 30 ítems

	30-1F	30-2F	30-3FD	30-3FE	30-4FD	30-4FE	30-5FD	30-5FE
Media	17.233	15.091	15.023	15.003	15.489	15.058	14.642	15.004
Desv. Tip.	8.399	6.941	8.269	5.721	8.117	5.153	7.998	4.652
Alfa	.956	.901	.941	.839	.937	.792	.935	.738
N.º de comp.	4	4	5	6	6	6	6	6
1.º autov. % var.	13.383 (44.6%)	7.981 (26.6%)	11.092 (37%)	5.736 (19.1%)	10.683 (35.6%)	5.071 (16.9%)	10.449 (34.8%)	3.845 (12.8%)
N.º de fact.	4	4	5	6	6	6	6	6
1.º autov. % var.	12.973 (43.2%)	7.566 (25.2%)	10.694 (35.6%)	5.329 (17.8%)	10.279 (34.3%)	4.655 (15.5%)	10.027 (33.4%)	3.435 (11.5%)

Tabla 3
Descripción de las bases de 50 ítems

	50-1F	50-2F	50-3FD	50-3FE	50-4FD	50-4FE	50-5FD	50-5FE
Media	20.446	23.235	23.409	23.667	23.876	23.335	22.538	23.661
Desv. Tip.	14.108	11.630	13.995	9.978	13.788	8.891	13.443	8.379
Alfa	.971	.944	.967	.914	.965	.889	.962	.871
N.º de comp.	4	4	5	5	5	5	5	5
1.º autov. % var.	22.661 (45.3%)	13.745 (27.5%)	19.221 (38.4%)	10.699 (21.3%)	18.479 (37%)	8.387 (16.8%)	17.286 (34.6%)	7.106 (14.2%)
N.º de fact.	4	4	5	5	5	5	5	5
1.º autov. % var.	22.257 (44.5%)	13.322 (26.6%)	18.859 (37.7%)	10.231 (20.5%)	18.118 (36.2%)	7.998 (16%)	17.664 (35.3%)	6.708 (13.4%)

resultado parece confirmar las críticas vertidas sobre α como indicador de unidimensionalidad (Green, Lissitz y Mulaik, 1977; Hattie, 1984, 1985). Indicar, finalmente, que el número de factores (y componentes) empíricos retenidos fue determinado de acuerdo al criterio de Kaiser (1960).

Sin duda, la pregunta central es ¿qué se está midiendo cuando se estiman las puntuaciones de los sujetos desde modelos unidimensionales cuando los datos son mani-

fiestamente multidimensionales?. Para tratar de dar respuesta a esta cuestión se ha calculado la matriz de correlaciones entre los diferentes tipos de puntuaciones calculadas (factoriales y θ).

Para una mayor claridad expositiva se irán comentando los ocho tipos distintos de tests simulados uno a uno. En cada una de las tablas que acompañan a estos comentarios se presenta la matriz de correlaciones.

En las muestras con un sólo factor, em-

Tabla 4
Matriz de correlaciones de la base 30-1F

	PAF1	PC1	P.Emp.	θ_1	θ_3
PAF1	1				
PC1	0.999	1			
P.Emp.	0.999	0.999	1		
θ_1	0.975	0.974	0.977	1	
θ_3	0.980	0.979	0.982	0.989	1

Tabla 5
Matriz de correlaciones de la base 50-1F

	PAF1	PC1	P.Emp.	θ_1	θ_3
PAF1	1				
PC1	0.999	1			
P.Emp.	0.999	0.999	1		
θ_1	0.979	0.979	0.985	1	
θ_3	0.982	0.982	0.988	0.999	1

pleadas como línea base, se observa (tablas 4 y 5) que, como era de esperar todos los tipos de puntuaciones correlacionan intensamente entre sí.

El primer caso multidimensional, bases 30-2F y 50-2F, presenta (tablas 6 y 7) unas correlaciones muy altas entre las puntuaciones factoriales alcanzadas tanto desde el análisis de componentes principales como desde el de factores principales, lo cual era esperable tras comprobar la similitud de los resultados alcanzados por uno y otro método mostrada en la descripción de las muestras llevada a cabo anteriormente. Este comportamiento se da en todas las muestras anali-

zadas, por lo que no se volverá a hacer referencia a él.

Las estimaciones realizadas desde el modelo logístico de un parámetro aparecen relacionadas tanto con las puntuaciones factoriales asociadas al primer factor como a las asociadas al segundo factor, en tanto que las estimaciones obtenidas desde el modelo de tres factores sólo muestran una elevada relación con el primer factor. También puede apreciarse que las correlaciones entre las estimaciones de θ y las puntuaciones empíricas son muy altas, especialmente en el caso de θ_1 . Otro aspecto que cabe destacar de lo presentado en este punto, como por otra

Tabla 6
Matriz de correlaciones de la base 30-2F

	PAF1	PAF2	PC1	PC2	P.Emp.	θ_1	θ_3
PAF1	1						
PAF2	0.000	1					
PC1	0.999	0.000	1				
PC2	0.000	0.999	0.000	1			
P.Emp.	0.730	0.683	0.730	0.683	1		
θ_1	0.722	0.679	0.722	0.680	0.991	1	
θ_3	0.892	0.289	0.892	0.290	0.845	0.847	1

Tabla 7
Matriz de correlaciones de la base 50-2F

	PAF1	PAF2	PC1	PC2	P.Emp.	θ_1	θ_3
PAF1	1						
PAF2	0.000	1					
PC1	0.999	0.002	1				
PC2	0.002	0.999	0.000	1			
P.Emp.	0.855	0.517	0.856	0.516	1		
θ_1	0.849	0.509	0.851	0.507	0.989	1	
θ_3	0.953	0.198	0.955	0.196	0.915	0.916	1

parte era de esperar teóricamente, es el comportamiento absolutamente paralelo que ofrecen las estimaciones realizadas a partir del modelo de Rasch y las puntuaciones empíricas.

Respecto al número de items cabe decir que básicamente se producen los mismos hechos en la base de 30 items y en la de 50 items, dándose únicamente leves diferencias en las cuantías de las correlaciones, de manera que en la base de 50 items parecen tomar valores un poco más extremos.

El siguiente modelo analizado correspon-

de al caso de tres factores con uno dominante, bases 30-3FD y 50-3FD, cuyos resultados pueden consultarse en las tablas 8 y 9. En este caso se da un sólo patrón compartido por las puntuaciones empíricas y las estimaciones hechas desde los dos modelos unidimensionales. Existe únicamente relación entre esas tres puntuaciones y el primer factor, al margen, claro está, de altas correlaciones entre ellas. En esta ocasión no se observan diferencias entre las dos bases.

En los tests simulados con cuatro y cinco factores similares a los referidos en el párra-

Tabla 8
Matriz de correlaciones de la base 30-3FD

	PAF1	PAF2	PAF3	PC1	PC2	PC3	P.Emp.	θ_1	θ_3
PAF1	1								
PAF2	-0.002	1							
PAF3	0.000	0.002	1						
PC1	0.999	0.001	0.000	1					
PC2	0.004	-0.999	0.027	0.000	1				
PC3	0.000	0.030	0.998	0.000	0.000	1			
P.Emp.	0.999	0.003	0.014	0.999	0.001	0.013	1		
θ_1	0.981	0.009	0.031	0.982	-0.007	0.031	0.982	1	
θ_3	0.981	0.018	-0.011	0.983	-0.017	-0.010	0.982	0.997	1

Tabla 9
Matriz de correlaciones de la base 50-3FD

	PAF1	PAF2	PAF3	PC1	PC2	PC3	P.Emp.	θ_1	θ_3
PAF1	1								
PAF2	-0.001	1							
PAF3	-0.001	0.000	1						
PC1	0.999	0.000	0.000	1					
PC2	-0.001	0.999	0.002	0.000	1				
PC3	-0.001	-0.002	0.999	0.000	0.000	1			
P.Emp.	0.999	-0.006	0.009	0.999	-0.007	0.008	1		
θ_1	0.984	0.012	0.053	0.985	0.012	0.053	0.984	1	
θ_3	0.968	-0.102	0.054	0.968	-0.102	0.055	0.968	0.975	1

fo anterior (30-4FD, 50-4FD, 30-5FD y 50-5FD) se dan las mismas pautas descritas para aquellos y por tanto todo lo dicho allí es aplicable punto por punto a estos dos tipos de bases (Tablas 10, 11, 12 y 13).

Las muestras 30-3FE y 50-3FE exhiben un funcionamiento general muy similar al mostrado por sus equivalentes de dos factores, aunque quizás convenga hacer algunas

leves matizaciones. En la muestra de 30 items, θ_1 se relaciona con los factores uno y tres, mientras que en la muestra de 50 items aunque las relaciones más intensas se dan con los factores uno y dos el tercer factor también es captado y se da con él una relación moderada. Por su parte θ_3 parece seguir relacionándose de manera principal con el primer factor. Sigue manteniéndose el para-

Tabla10
Matriz de correlaciones de la base 30-4FD

	PAF1	PAF2	PAF3	PAF4	PC1	PC2	PC3	PC4	P.Emp	θ_1	θ_3
PAF1	1										
PAF2	-.003	1									
PAF3	.002	.004	1								
PAF4	.001	-.002	.001	1							
PC1	.999	.001	-.001	.000	1						
PC2	-.004	.997	-.024	.006	.000	1					
PC3	.003	.029	.997	-.004	.000	.000	1				
PC4	.001	-.008	.005	.997	.000	.000	.000	1			
P.Emp	.999	.018	.001	.008	.999	.017	.002	.009	1		
θ_1	.981	.038	.011	.017	.982	.039	.012	.018	.983	1	
θ_3	.982	-.017	.003	.000	.984	-.017	.002	.001	.983	.995	1

Tabla11
Matriz de correlaciones de la base 50-4FD

	PAF1	PAF2	PAF3	PAF4	PC1	PC2	PC3	PC4	P.Emp	θ_1	θ_3
PAF1	1										
PAF2	.001	1									
PAF3	-.002	.004	1								
PAF4	.000	-.001	.000	1							
PC1	.999	.001	-.001	.000	1						
PC2	.002	.999	-.027	.004	.000	1					
PC3	-.002	.031	.999	-.000	.000	.000	1				
PC4	-.001	-.005	.001	.999	.000	.000	.000	1			
P.Emp	.999	-.009	-.005	-.005	.999	-.008	-.006	-.005	1		
θ_1	.986	.011	.035	.024	.987	-.012	.036	.025	.986	1	
θ_3	.981	-.069	.072	.007	.982	-.071	.071	.008	.981	.989	1

Tabla 12
Matriz de correlaciones de la base 30-5FD

	PAF1	PAF2	PAF3	PAF4	PAF5	PC1	PC2	PC3	PC4	PC5	P.Emp	θ_1	θ_3
PAF1	1												
PAF2	.001	1											
PAF3	-.006	-.010	1										
PAF4	-.001	-.002	.004	1									
PAF5	-.005	-.006	-.004	-.012	1								
PC1	.999	.000	.001	.000	.001	1							
PC2	-.001	.964	.230	.009	.044	.000	1						
PC3	-.007	-.234	.968	-.019	.015	.000	.000	1					
PC4	-.006	-.052	-.017	.446	.875	.000	.000	.000	1				
PC5	-.002	-.008	-.030	-.890	.456	.000	.000	.000	.000	1			
P.Emp	.999	.004	.008	-.002	.006	.999	.005	.006	.002	.004	1		
θ_1	.983	.004	.016	-.001	.025	.986	.011	.015	.025	.012	.986	1	
θ_3	.985	-.016	-.010	-.005	.008	.987	-.013	-.006	.011	.010	.987	.996	1

Tabla 13
Matriz de correlaciones de la base 50-5FD

	PAF1	PAF2	PAF3	PAF4	PAF5	PC1	PC2	PC3	PC4	PC5	P.Emp	θ_1	θ_3
PAF1	1												
PAF2	.000	1											
PAF3	-.002	.004	1										
PAF4	-.002	.003	.000	1									
PAF5	-.003	.002	-.007	-.003	1								
PC1	.999	.000	.001	.000	.001	1							
PC2	.000	.993	-.086	-.024	-.013	.000	1						
PC3	-.004	.092	.993	.005	.053	.000	.000	1					
PC4	-.002	.029	-.012	.997	.063	.000	.000	.000	1				
PC5	-.004	.008	-.062	-.067	.995	.000	.000	.000	.000	1			
P.Emp	.999	.017	-.014	.002	.000	.999	.018	-.013	.002	.000	1		
θ_1	.986	.039	.006	.016	.041	.987	.039	.012	.019	.041	.987	1	
θ_3	.973	.133	.018	-.004	.067	.975	.133	.034	.002	.069	.976	.983	1

Tabla 14
Matriz de correlaciones de la base 30-3FE

	PAF1	PAF2	PAF3	PC1	PC2	PC3	P.Emp.	θ_1	θ_3
PAF1	1								
PAF2	0.002	1							
PAF3	0.000	-0.002	1						
PC1	0.999	-0.045	-0.002	1					
PC2	0.045	0.999	0.009	0.000	1				
PC3	0.002	-0.011	0.999	0.000	0.000	1			
P.Emp.	0.828	0.027	0.559	0.826	0.070	0.560	1		
θ_1	0.824	0.030	0.560	0.821	0.073	0.561	0.997	1	
θ_3	0.830	-0.377	0.236	0.845	-0.337	0.241	0.807	0.808	1

Tabla 15
Matriz de correlaciones de la base 50-3FE

	PAF1	PAF2	PAF3	PC1	PC2	PC3	P.Emp.	θ_1	θ_3
PAF1	1								
PAF2	0.000	1							
PAF3	0.000	0.000	1						
PC1	0.999	-0.001	-0.001	1					
PC2	0.001	0.999	-0.008	0.000	1				
PC3	0.001	0.009	0.999	0.000	0.000	1			
P.Emp.	0.574	0.718	0.394	0.573	0.715	0.401	1		
θ_1	0.559	0.718	0.401	0.558	0.715	0.408	0.994	1	
θ_3	0.830	0.305	0.336	0.829	0.304	0.340	0.827	0.835	1

lelismo entre θ_1 y las puntuaciones empíricas (véanse tablas 14 y 15).

Al examinar (tablas 16 y 17) las bases de cuatro factores equivalentes (30-4FE y 50-4FE) se encuentra con que en el caso de 30 items, θ_1 y las puntuaciones empíricas parecen captar principalmente dos de los factores, con una cierta presencia de un tercer factor y obviando totalmente el cuarto factor. En la muestra de 50 items, aun cuando pare-

ce tener un poco más de peso el factor uno, son captados los cuatro factores. Este comportamiento puede considerarse como paralelo a lo mencionado más arriba al describir los resultados en las correspondientes bases de tres factores. Por su parte θ_3 sigue manteniendo su estrecha relación con el primer factor.

Obsérvense, por último, los casos de cinco factores de igual fuerza. La base 30-5FE

MODELOS UNIDIMENSIONALES CON DATOS MULTIFACTORIALES

Tabla 16
Matriz de correlaciones de la base 30-4FE

	PAF1	PAF2	PAF3	PAF4	PC1	PC2	PC3	PC4	P.Emp	θ_1	θ_3
PAF1	1										
PAF2	-.001	1									
PAF3	.000	.007	1								
PAF4	.001	-.003	.000	1							
PC1	.999	.006	.002	-.003	1						
PC2	-.007	.984	-.164	.016	.000	1					
PC3	-.005	.173	.984	-.005	.000	.000	1				
PC4	.005	-.019	.009	.997	.000	.000	.000	1			
P.Emp	.507	.341	.787	.079	.513	.198	.831	.080	1		
θ_1	.504	.342	.786	.079	.509	.200	.831	.081	.999	1	
θ_3	.916	.244	.173	.049	.921	.208	.208	.052	.690	.693	1

Tabla 17
Matriz de correlaciones de la base 50-4FE

	PAF1	PAF2	PAF3	PAF4	PC1	PC2	PC3	PC4	P.Emp	θ_1	θ_3
PAF1	1										
PAF2	.000	1									
PAF3	.000	.000	1								
PAF4	-.001	-.003	.000	1							
PC1	.999	.004	.000	.002	1						
PC2	-.005	.999	-.001	.017	.000	1					
PC3	.001	.002	.999	.004	.000	.000	1				
PC4	-.003	-.021	-.004	.999	.000	.000	.000	1			
P.Emp	.761	.455	.343	.306	.764	.457	.345	.295	1		
θ_1	.754	.454	.349	.311	.757	.456	.351	.300	.998	1	
θ_3	.919	-.028	.209	.118	.919	-.030	.211	.116	.795	.798	1

Tabla 18
Matriz de correlaciones de la base 30-5FE

	PAF1	PAF2	PAF3	PAF4	PAF5	PC1	PC2	PC3	PC4	PC5	P.Emp	θ_1	θ_3
PAF1	1												
PAF2	-.001	1											
PAF3	-.001	-.002	1										
PAF4	-.001	.001	-.003	1									
PAF5	.000	.000	.000	.000	1								
PC1	.997	.007	.018	.007	.001	1							
PC2	-.009	.996	.045	-.021	.000	.000	1						
PC3	-.020	-.046	.994	.071	.003	.000	.000	1					
PC4	-.006	.026	-.075	.996	.003	.000	.000	.000	1				
PC5	-.001	.001	-.003	-.004	.994	.000	.000	.000	.000	1			
P.Emp	.611	.616	.050	.279	.398	.622	.607	.035	.287	.398	1		
θ_1	.609	.615	.050	.279	.401	.620	.607	.035	.287	.401	.999	1	
θ_3	.943	.145	.087	-.001	.059	.942	.142	.062	-.008	.059	.694	.695	1

ofrece resultados semejantes a los comentarios en los casos de tres y cuatro factores (tabla 18). En su relación con θ_1 y las puntuaciones empíricas aunque el factor uno y dos presentan mayor fuerza también el cuatro y el cinco están presentes, siendo el tres ignorado. θ_3 al igual que en las bases restantes correlaciona únicamente con las puntuaciones factoriales del primer factor. Por su parte los datos de 50-5FE no se ajustan de manera tan clara al patrón que hasta ahora se había venido especificando en las bases con sus mismas características. Como puede verse en la tabla 19 el comportamiento que en este caso presentan θ_1 y las puntuaciones empíricas se asemeja bastante al de θ_3 , aun cuando se atisbe una muy leve tendencia a mostrar alguna relación con los factores dos, tres y cinco. Parece, por lo tanto, que en esta ocasión θ_1 se ve atrapado por el

primer factor al igual que ocurría habitualmente con θ_3 .

Resumiendo lo dicho hasta ahora, y coincidiendo básicamente con lo apuntado por Reckase, cuando se trabaja con factores independientes el modelo de tres parámetros estima únicamente un factor mientras que el modelo de un parámetro parece estimar una combinación de todos los factores. Cuando hay un factor dominante el comportamiento de los dos modelos es el mismo estimando ambos ese factor preponderante. También resulta interesante la confirmación empírica de las puntuaciones empíricas como estimador suficiente de θ en el modelo de un parámetro, como se puede derivar del comportamiento absolutamente paralelo de las puntuaciones empíricas y θ_1 , así como de las altísimas correlaciones entre ellas, que se encuentran en las diferentes bases de datos.

Tabla 19
Matriz de correlaciones de la base 50-5FE

	PAF1	PAF2	PAF3	PAF4	PAF5	PC1	PC2	PC3	PC4	PC5	P.Emp	θ_1	θ_3
PAF1	1												
PAF2	.001	1											
PAF3	-.003	.001	1										
PAF4	-.001	.002	.000	1									
PAF5	-.002	-.001	-.002	.001	1								
PC1	.999	-.013	.016	.002	.007	1							
PC2	.014	.999	-.010	-.014	-.003	.000	1						
PC3	-.019	.012	.999	.006	.016	.000	.000	1					
PC4	-.002	.016	-.007	.999	-.011	.000	.000	.000	1				
PC5	-.009	-.004	-.019	.012	.998	.000	.000	.000	.000	1			
P.Emp	.897	.272	.273	.021	.209	.900	.281	.266	.019	.198	1		
θ_1	.893	.268	.278	.026	.215	.896	.277	.270	.024	.204	.998	1	
θ_3	.859	-.074	.381	.019	-.034	.867	-.068	.367	.016	-.045	.850	.850	1

CONCLUSIONES GENERALES

En última instancia, el objetivo fundamental de los trabajos como el aquí desarrollado es tratar de dar una respuesta a la pregunta ¿qué se está midiendo cuando se emplean modelos unidimensionales sobre datos con varios factores?, o desde un punto de vista aplicado ¿son válidas las estimaciones de las puntuaciones de los sujetos hechas mediante modelos unidimensionales sobre datos multifactoriales?. Por lo encontrado aquí, y en concordancia con lo presentado por Reckase y otros autores, cuando entre los diversos factores existe uno claramente dominante parece claro que los modelos unidimensionales tienden a medir ese factor dominante. Cuando los diferentes factores que subyacen a los datos poseen pesos similares, el comportamiento de los modelos

varía. El modelo de un parámetro recoge una combinación de todos los factores presentes, en tanto que el modelo de tres parámetros únicamente capta el primer factor.

El mensaje a transmitir al usuario sería que, evidentemente, la violación del supuesto de unidimensionalidad influye en la estimación de las puntuaciones de los sujetos, sin embargo, cuando en los datos haya un factor claramente dominante las estimaciones pueden ser empleadas con cierta garantía para medir ese factor. Si no existe factor dominante el asunto es más delicado y probablemente lo más práctico, y sensato, en casos de bases multifactoriales con factores de igual o parecida potencia, sea tratar de construir varios subtests de factor único y emplearlos individualmente. En cuanto a la preferencia de los usuarios por el modelo de un parámetro debido a su sencillez, parece bas-

tante robusto a violaciones no extremas de la unidimensionalidad, si bien el esquema utilizado aquí no es en puridad el ideal para investigar este aspecto, habría que utilizar estimaciones mediante modelos multidimensionales, que es la línea de futuro en la que pretendemos trabajar.

REFERENCIAS

- Ackerman, T. A. (1989). Unidimensional IRT calibration of compensatory and noncompensatory multidimensional items. *Applied Psychological Measurement, 13* (2), 113-127.
- Ansley, T. N. y Forsyth, R. A. (1985). An examination of the characteristics of unidimensional IRT parameter estimates derived from two-dimensional data. *Applied Psychological Measurement, 9* (1), 37-48.
- Berstein, I. H. (1988). *Applied multivariate analysis*. New York: Springer-Verlag.
- Doody-Bogan, E. N. y Yen, W. M. (1983). *Detecting multidimensionality and examining the effects of vertical equating with the three parameter logistic model*. Comunicación presentada en la reunión anual de la AERA, Montreal.
- Drasgow, F. y Parsons, C. K. (1983). Application of unidimensional item response theory model to multidimensional data. *Applied Psychological Measurement, 7* (2), 189-199.
- Green, S. B.; Lissitz, R. W. y Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement, 37*, 827-838.
- Hambleton, R. K. y Swaminathan, H. (1985). *Item response theory. Principles and applications*. Boston, MA: Kluwer-Nijhoff Publishers.
- Harrison, D. A. (1986). Robustness of IRT parameter estimation to violations of unidimensionality assumption. *Journal of Educational Statistics, 11* (2), 95-115.
- Hattie, J. (1984). An empirical study of various indices for determining unidimensionality. *Multivariate Behavioral Research, 19*, 49-78.
- Hattie, J. (1985). Methodology review: assessing unidimensionality of tests and items. *Applied Psychological Measurement, 9* (2), 139-164.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.
- Muñoz, J. (1990). *Teoría de respuesta a los items*. Madrid: Pirámide.
- Muñoz, J.; Rogers, J. y Swaminathan, H. (1989). Robustez de las estimaciones del modelo de Rasch en presencia de aciertos al azar y discriminación variable de los items. *Anuario de Psicología, 4* (3), 83-97.
- Reckase M. D. (1979). Unifactor latent trait models applied to multifactor tests: results and implications. *Journal of Educational Statistics, 4* (3), 207-230.
- Way, W. D.; Ansley, T.N. y Forsyth, R. A. (1988). The comparative effects of compensatory and noncompensatory two-dimensional data on unidimensional IRT estimates. *Applied Psychological Measurement, 12* (3), 239-259.
- Wherry, R. J. Sr.; Naylor, J. C.; Wherry, R. J. Jr. y Fallis, R. F. (1965). Generating multiple samples of multivariate data with arbitrary population parameters. *Psychometrika, 30*, 303-313.
- Yen, W. M. (1984). Effects of local dependence on the fit and equating performance of the three parameter logistic model. *Applied Psychological Measurement, 8* (2), 125-145.

Accepted el 11 de diciembre de 1993