

Aplicación de un modelo politómico de TRI para la obtención de medidas conmensurables en una escala de ajuste persona-organización.

M. Carmen Ximénez y Javier Revuelta*

Universidad Autónoma de Madrid

En este artículo se aborda el problema del uso de métodos directos para la comprobación del criterio de 'conmensurabilidad de unidades de medida' en las medidas sobre *ajuste persona-ambiente* mediante la aplicación de un modelo politómico de rasgo latente. Se utilizó el modelo de respuesta nominal de Bock (1972; 1997) para evaluar el ajuste estadístico de las puntuaciones en las escalas de medida P y O de un cuestionario de ajuste persona-organización con datos reales de 591 sujetos. Los resultados indicaron que el método propuesto constituye una aproximación inicial al problema apropiada, ya que proporciona puntuaciones normalizadas lo que permite una interpretación más sencilla de las unidades de medida en P y O; aunque también presenta una serie de limitaciones que han de ser consideradas en futuras investigaciones.

Palabras clave: Teoría de Respuesta a los Ítems (TRI), Modelos politómicos de rasgo latente, conmensurabilidad en unidades de medida, ajuste persona-organización (*ajuste P-O*).

Uno de los problemas metodológicos que presentan las medidas de ajuste persona-ambiente en general, y las de ajuste persona-organización (o ajuste P-O) en particular, consiste en la obtención de mediciones para un mismo sujeto en dos dominios teóricos diferentes: de un lado el de sus propias necesidades y motivaciones; y de otro, el del grado en que éstas se ven satisfechas en su organización de trabajo (véase Ximénez, 1998). Para poder evaluar el grado de ajuste o correspondencia de la persona (en lo sucesivo P) a la organización (O) es necesario obtener y comparar medidas

* Queremos agradecer a Rafael San Martín por sus contribuciones a este trabajo. Esta investigación ha sido financiada, en parte, por los proyectos SEC93-0801 y PB97-0049 de la DGICYT. La correspondencia en referencia a este artículo ha de enviarse a: Javier Revuelta y Carmen Ximénez; Departamento de Psicología Social y Metodología; Universidad Autónoma de Madrid; Ciudad Universitaria de Cantoblanco. 28049 MADRID; Tfno. (91) 397 52 03. Fax: (91) 397 52 15; e.mail: javier.revuelta@uam.es

independientes de ambos constructos; y que dichas medidas cumplan ciertos requisitos para que la comparación entre ellas tenga sentido. Uno de los más importantes y discutidos es que sean conmensurables o que se expresen en dominios teóricos paralelos. Rounds, Dawis y Lofquist (1987) establecieron tres criterios para el cumplimiento del requisito de la conmensurabilidad:

1.- *Conmensurabilidad de conceptos*. Significa que las medidas de P y O deben reflejar aspectos teóricos equivalentes. El procedimiento habitualmente empleado para garantizar este criterio consiste en utilizar los mismos ítems en P y O, no habiéndose empleado un método directo para su comprobación.

2.- *Conmensurabilidad de unidades de medida*. Se refiere a que las unidades o puntos de anclaje empleados en la escala de medida de P sean equivalentes a los empleados en la de O. El cumplimiento de este supuesto es difícil ya que implica que el incremento o descenso de una unidad en la escala P sea igual (en el mejor de los casos) o comparable (en el peor de los casos) al mismo incremento o descenso en la escala O. Rounds et al. (1987) señalaban que este criterio puede evaluarse de forma indirecta mediante técnicas de escalamiento pero, al igual que el anterior, no ha sido demostrado con métodos directos. Aún no se ha propuesto una solución satisfactoria a este problema y esto se debe en parte, como subraya Edwards (1994), a que en la práctica se ha resuelto utilizando escalas de medida con categorías de respuestas equivalentes en P y O; sin embargo esto no garantiza que su sentido sea el mismo en ambas escalas.

3.- *Conmensurabilidad de estructura*. Consiste en la organización paralela y equivalente tanto de las características incluidas en las medidas de P como en las de O. Este criterio se evalúa comprobando si ambas medidas obtienen estructuras factoriales similares mediante análisis factoriales separados en los ítems de P y O, y ha sido el más investigado de los tres (para más información véase Ximénez y San Martín, 1998).

El sentido de la conmensurabilidad de unidades de medida, puede no ser evidente a primera vista y confundirse con la similaridad numérica de las escalas resultantes. Dos variables pueden estar gobernadas por distribuciones idénticas y no ser conmensurables. Por ejemplo, la estatura en centímetros en un grupo de personas y la concentración de un material en muestras de excavaciones geológicas pueden seguir una distribución normal (1.75, 6); sin embargo, cualquier comparación entre dichas variables carece de sentido.

Para que dos medidas sean conmensurables deben obtenerse sobre la base de puntos de referencia comunes, aunque los parámetros de sus distribuciones (e incluso la propia forma de la distribución) sean diferentes. Por ejemplo, si sabemos que la temperatura media anual en Atenas (Grecia) es de 17° C. y en Okhotsk (CEI) de -6° C., ambas medidas (o una distribución de medidas que pudieran obtenerse) son directamente comparables a pesar de su gran diferencia.

Este artículo se centra específicamente en el problema de la satisfacción del criterio de la conmensurabilidad de unidades de medida, uno de los

menos considerados en la investigación previa pero sin embargo más citado en las revisiones sobre problemas en las medidas de ajuste persona-ambiente (ver p.e. Rounds, et al., 1987). Este objetivo se aborda mediante la aplicación de un modelo de rasgo latente (ver Hambleton y Van der Linden, 1997, para una revisión) en una medida de ajuste persona-organización con datos reales de 591 sujetos. El cuestionario empleado consiste en 30 ítems divididos en dos grupos de 15, que evalúan las necesidades de la persona (P) y el grado en que ésta las ve promovidas por su organización (O), respectivamente. Los ítems pueden considerarse gemelos, en el sentido de que existe una correspondencia 1 a 1 entre cada uno de ellos en P y en O. Cada pareja de ítems tiene exactamente la misma redacción, salvo que en un caso el contenido se refiere a P y en otro a O. Por ejemplo:

Para mi ORGANIZACIÓN es importante que... Nada Poco Medio Bastante Mucho 1 2 3 4 5	...Yo obtenga un sueldo competitivo ...	PARA MÍ es importante que... Nada Poco Medio Bastante Mucho 1 2 3 4 5
---	---	---

Nota: Véase Ximénez y San Martín (1998; pág. 257) para consultar cuestionario completo.

En este artículo se presenta un método para el análisis de la conmensurabilidad de unidades de medida en un cuestionario de ajuste persona-ambiente (Ximénez y San Martín, 1998). El método ha sido diseñado para obtener medidas conmensurables en P y en O, lo que no significa que deban ser parecidas ni individual ni grupalmente. El método se basa en un modelo de rasgo latente que se describe a continuación. En el siguiente apartado se indica la forma de adaptar dicho modelo al problema planteado. Finalmente, el artículo incluye un ejemplo de aplicación de este método a una muestra de datos reales.

El modelo.

Los datos recogidos son de naturaleza politómica, es decir la respuesta a cada ítem se expresa mediante una de las 5 alternativas de respuesta. En adelante se utilizará el símbolo y para referirse a la matriz de datos del experimento, donde cada uno de sus elementos y_{si} es la respuesta del sujeto s al ítem i . El número total de sujetos se indica por S y el de ítems por I .

En la literatura aparecen descritos distintos modelos de rasgo latente para datos politómicos. En este trabajo el objetivo se centra en obtener las puntuaciones de las personas en las escalas conmensurables P y O, no tanto en conocer las propiedades de los ítems en sí. Por esta razón el modelo escogido es el de respuesta nominal (MRN) propuesto por Bock (1972; 1997) que es el de mayor generalidad. Otros modelos muy conocidos, como el modelo de crédito parcial (Masters, 1982) son casos particulares del MRN,

por lo que su ajuste estadístico a unos datos reales es por definición igual o peor que el obtenido con el modelo general.

El MRN relaciona la probabilidad de que el sujeto escoja cada una de las C alternativas de respuesta a un ítem en función de su nivel de rasgo θ_i y de los parámetros de la alternativa α_a y β_a . La probabilidad de que el sujeto escoja la alternativa a se define como:

$$P(y_{si} = a | \theta_i, \varepsilon_i) = \frac{\exp(\alpha_a + \beta_a \theta_i)}{\sum_{c=1}^C \exp(\alpha_c + \beta_c \theta_i)} \quad (1)$$

Donde ε_i es el vector de parámetros del ítem i , que contiene los valores de α_c y β_c de todas sus alternativas. Si se asume independencia local, la probabilidad de que un sujeto s obtenga un vector de respuestas y_s es:

$$f(y_s | \theta, \varepsilon) = \prod_{i=1}^I P(y_{si} | \theta_i, \varepsilon_i)$$

Donde el vector ε recoge los parámetros de todos los ítems. En el problema abordado en este artículo se intenta medir los niveles de rasgo de las necesidades de la persona: θ_p , y del grado en que ésta las ve promovidas en su organización: θ_o en cada sujeto, utilizando una escala común para ambos.

En el cuestionario empleado los ítems se distribuyen por parejas para P y O, por lo que la matriz de datos completos está formada por las submatrices y_p e y_o . Además es posible definir las funciones $f(y_{sp} | \theta_p, \varepsilon_p)$ y $f(y_{so} | \theta_o, \varepsilon_o)$ que indican la probabilidad de que un sujeto obtenga un vector de respuestas determinado en los ítems de cada mitad. Estas funciones dependen de los dos niveles de rasgo de la persona y de los parámetros de los ítems, que en principio pueden ser diferentes para P y para O. Al hacer esta definición se está asumiendo que las respuestas dadas a una escala dependen únicamente del rasgo medido por esa escala. Es decir, que el valor θ_p de un sujeto no afecta a sus respuestas en O y viceversa.

Debido a la arbitrariedad de la variable latente es necesario imponer alguna restricción en los parámetros para identificar el modelo. En la literatura aparecen descritas varias de ellas sin que existan criterios para escoger una u otra. Por esta razón la restricción impuesta consiste en fijar a 0 los parámetros α_1 y β_1 de la primera alternativa de todos los ítems.

Tal y como está planteado el problema es posible obtener estimaciones separadas para ambas escalas, sin embargo dichas estimaciones no pueden considerarse commensurables porque no existe ningún punto de anclaje entre

ambas. La obtención de mediciones commensurables en ambos niveles de rasgo puede hacerse en base a puntos de referencia comunes. Este punto de referencia común consiste en asumir los mismos parámetros para cada uno de los ítems gemelos en P y en O; es decir, se impone la restricción $\epsilon_p = \epsilon_o$. El sentido de esta restricción es que un mismo valor de θ_p y θ_o produce las mismas probabilidades esperadas de respuesta en ambas escalas. En la estimación de parámetros este proceso se invierte, y a partir de las probabilidades observadas se obtiene el valor de rasgo correspondiente.

Estimación de parámetros.

Para estimar los parámetros del modelo se utiliza el método marginal, en el cual se asume que los valores de rasgo constituyen una variable aleatoria no observada. Supongamos que θ denota el conjunto de niveles de rasgo involucrados en el test. El vector θ se compone entonces de los dos valores θ_p y θ_o del total de sujetos. En lo sucesivo se considera que cada elemento de θ es una variable aleatoria distribuida según la normal estandarizada $g(\cdot)$.

Asumir que cada sujeto se define por dos niveles independientes de rasgo, que los parámetros de los ítems son iguales en P y O, y que existe independencia estocástica entre y_p e y_o , es equivalente a suponer que las respuestas a cada par de ítems gemelos son respuestas al mismo ítem dadas por dos sujetos diferentes. De esta forma puede considerarse que el número de sujetos es igual a 2 veces el número de sujetos reales y que el número de ítems es igual al número de pares de ítems gemelos. Con este diseño la probabilidad de los datos completos es:

$$f(y | \theta, \epsilon) = \prod_{s=1}^S \prod_{i=1}^I P(y_{si} | \theta_s, \epsilon_i) g(\theta_s) \quad (2)$$

Donde S es igual a $2 \times n^\circ$ de sujetos reales, I es el número de pares de ítems y ϵ_i son los parámetros del par i . En primer lugar es necesario estimar ϵ . Como θ es una variable no observada la expresión (2) no puede evaluarse directamente, y es necesario calcular la función de probabilidad de la matriz y marginalizando la función de densidad de los datos completos:

$$f(y | \epsilon) = \int \prod_{s=1}^S \prod_{i=1}^I P(y_{si} | \theta_s, \epsilon_i) g(\theta_s) d\theta \quad (3)$$

El estimador de ϵ se obtiene como aquel valor que maximiza la probabilidad marginal (3), en lo que se conoce como método de máxima verosimilitud marginal (Bock y Aitkin, 1981). La forma de llevarlo a cabo

consiste en aplicar el algoritmo EM (Dempster, Laird y Rubin, 1977) implementado en el programa MULTLOG (Thissen, 1991).

A continuación, se obtienen las estimaciones de cada elemento de θ condicionadas en el valor estimado de ϵ . Para cada sujeto, dicha estimación se obtiene como el valor que maximiza la función de densidad posterior:

$$f(\theta | y_s, \epsilon) = \prod_{i=1}^I P(y_{si} | \theta, \epsilon_i) g(\theta)$$

Cabe esperar que los valores de θ así estimados se distribuyan aproximadamente según la normal estandarizada. El último paso del método consiste simplemente en dividir el vector θ en sus dos componentes, correspondientes a los valores de μ_p y μ_o de cada sujeto. Resulta muy sencillo interpretar μ_p y μ_o sabiendo que el conjunto de todas las θ se distribuye aproximadamente según el modelo normal estandarizado.

A continuación se comenta un estudio empírico en el que se aplica el método propuesto para satisfacer la conmensurabilidad de unidades de medida en el cuestionario de ajuste P-O descrito en la introducción de este artículo.

MÉTODO

Muestra y recogida de datos. El cuestionario se aplicó a una muestra de 591 sujetos, de los cuales 295 eran mujeres y 296 varones. Como el cuestionario se compone de 15 ítems expresados en dos escalas, el tamaño de la muestra es $591 \cdot 2 = 1.182$ sujetos. La recogida de datos se llevó a cabo mediante un envío por correo postal del cuestionario, siendo opcional y confidencial su contestación.

Instrumento de medida. Consistió en un cuestionario de tipo auto-informe compuesto de quince ítems presentados en el centro de una tabla de tres columnas. Primero se pedía a los sujetos que contestaran en la columna derecha indicando la importancia de cada uno de los aspectos incluidos en los ítems para sí mismo/a (escala P). Más tarde se pedía que volvieran al principio y contestaran en la columna izquierda indicando si creían que cada uno de los mismos aspectos era importante o no para su organización (escala O). El sujeto disponía de una escala tipo Likert de 5 puntos (siendo 1: nada importante; y 5: muy importante) para responder ambas medidas (para más detalles sobre la muestra, el procedimiento y el instrumento de medida ver Ximénez y San Martín, 1998, p. 240-241).

Análisis de datos. En primer lugar se llevaron a cabo análisis factoriales exploratorios separados sobre los ítems correspondientes a P y O. Se utilizó el método de extracción de factores de mínimos cuadrados generalizados fijado a un factor para examinar el supuesto de la unidimensionalidad. Asimismo, se obtuvieron otros indicadores para comprobar dicho supuesto en ambos grupos de ítems, como el cociente entre el primer autovalor y el segundo y el '*Scree test*' o la comparación gráfica de los autovalores obtenidos en los ítems de P y O. En segundo lugar se obtuvieron las estimaciones de ϵ y θ por el método propuesto. Para cada sujeto se obtuvieron sus valores μ_p y μ_o y se comparó la distribución de ambos en la muestra total. Todos los análisis se realizaron mediante los programas SPSS 7.5.2S. y MULTILOG.

RESULTADOS

La distribución de frecuencias de las respuestas observadas muestra que en la escala P la mayor parte de los sujetos escogen la categoría de respuesta más alta (número 5), que indica que conceden un gran valor a alguna de las características de la organización en que trabajan. En cambio, en la escala O las alternativas más populares son las de valor medio, que indican solamente un refuerzo moderado por parte de la organización. Además los datos muestran una ordenación de las respuestas, en el sentido de que existe una relación entre la frecuencia de elección y el número de alternativa. En la tabla 1 puede verse el número de sujetos que escoge cada una de las 5 alternativas de los 15 ítems de P y de O (para consultar resultados descriptivos sobre los ítems en P y O ver Ximénez y San Martín, 1998, p. 245-247).

Debido al formato del cuestionario, existe el peligro de que se produzcan dependencias entre las respuestas de los sujetos a los ítems de P y O. Dado que la independencia es un supuesto fundamental del método de estimación, su violación puede invalidar los estimadores de los parámetros. En la literatura aparecen descritos diversos estadísticos para comprobar la independencia local (Chen y Thissen, 1997), sin embargo no son aplicables a estos datos debido a que la distribución de los mismos sólo se conoce de forma asintótica y para utilizarlos con garantías es necesario que el tamaño muestral exceda ampliamente del número de posibles patrones de respuesta. Una primera aproximación al problema de la independencia local aparece descrita en Ximénez y San Martín (1998), donde se calculan las correlaciones entre los ítems de cada par en P y O. El resultado de este análisis es que para la mayoría de los pares la correlación no es significativa, por lo que de forma provisional se sugiere que al menos no existe relación lineal entre ambas respuestas.

Tabla 1:

Distribución de frecuencias de las respuestas observadas en las escalas P y O

ITEM	<i>Escala P</i>					<i>Escala O</i>				
	Alt. 1	Alt. 2	Alt. 3	Alt. 4	Alt. 5	Alt. 1	Alt. 2	Alt. 3	Alt. 4	Alt. 5
Item 1	1	16	99	322	153	103	208	188	76	16
Item 2	1	12	46	221	311	82	161	176	124	48
Item 3	9	25	87	257	213	86	150	193	116	46
Item 4	4	24	123	274	166	47	114	160	185	85
Item 5	1	3	49	256	282	63	143	164	159	62
Item 6	2	6	38	266	279	81	176	179	109	46
Item 7	1	0	55	257	278	65	140	170	146	70
Item 8	8	26	116	243	198	127	198	163	72	31
Item 9	2	14	110	277	188	74	146	176	121	74
Item 10	15	39	197	260	80	35	95	121	162	178
Item 11	15	67	214	227	68	133	180	134	96	48
Item 12	0	1	42	227	321	19	52	88	149	283
Item 13	2	24	71	217	277	153	186	175	59	18
Item 14	12	57	175	225	122	137	204	158	61	31
Item 15	15	71	203	216	86	140	208	175	55	13

Los resultados de los análisis factoriales exploratorios separados en los ítems correspondientes tanto a P como a O aparecen en la tabla 2. Como se observa, el primer factor explica el mayor porcentaje de varianza en ambos casos: el 23.32% en P y el 42.64% en O. Según el criterio de Carmines y Zeller (1979) estos resultados indican que solamente puede considerarse que existe unidimensionalidad en O, ya que se necesita que como mínimo el primer factor dé cuenta del 40% de la varianza. Sin embargo, según el criterio de Reckase (1979), que reduce la cifra al 20%, puede considerarse que ambos tests son unidimensionales. En cuanto a los restantes indicadores, el cociente entre el primer y segundo autovalor propuesto por Lumsden (1957; 1961) es 2.29 para P y 4.81 para O; lo que de nuevo indica que el supuesto de unidimensionalidad es más aceptable en O que en P.

Tabla 2:

Análisis de la unidimensionalidad en los ítems de las medidas de P y O

ITEM	Escala P		Escala O	
	Factor 1	h^2	Factor 1	h^2
Ítem 1	.267	.195	.621	.430
Ítem 2	.227	.148	.493	.290
Ítem 3	.408	.201	.670	.437
Ítem 4	.459	.288	.606	.417
Ítem 5	.464	.291	.649	.453
Ítem 6	.575	.411	.738	.577
Ítem 7	.519	.319	.682	.498
Ítem 8	.348	.222	.611	.407
Ítem 9	.432	.275	.565	.365
Ítem 10	.310	.105	.336	.267
Ítem 11	.282	.099	.473	.254
Ítem 12	.469	.217	.484	.374
Ítem 13	.466	.286	.748	.554
Ítem 14	.469	.378	.710	.513
Ítem 15	.499	.376	.787	.610
Autovalor, λ_i	1 ^{er} factor: 3.498 2 ^o factor: 1.53 Cociente $\lambda_1/\lambda_2 = 2.29$		1 ^{er} factor: 6.396 2 ^o factor: 1.331 Cociente $\lambda_1/\lambda_2 = 4.81$	
% Varianza	1 ^{er} factor: 23.32 % 2 ^o factor: 10.23 %		1 ^{er} factor: 42.64 % 2 ^o factor: 8.871 %	

Nota: Los análisis factoriales exploratorios se realizaron con método de extracción de mínimos cuadrados generalizados. En ambos casos se extrajo 1 solo factor tras 5 iteraciones. h^2 se refiere a las comunalidades.

La comparación entre los autovalores obtenidos en las soluciones factoriales en P y O puede verse en la figura 1. Como se observa, en ambos casos hay un salto claro entre el primer autovalor y el segundo, quedando por encima de los valores mínimos los de los dos primeros factores. Tomados en conjunto los resultados indican que el test satisface sólo moderadamente la hipótesis de unidimensionalidad (Hambleton y Swaminathan, 1985) y son notablemente más favorables para la escala O.

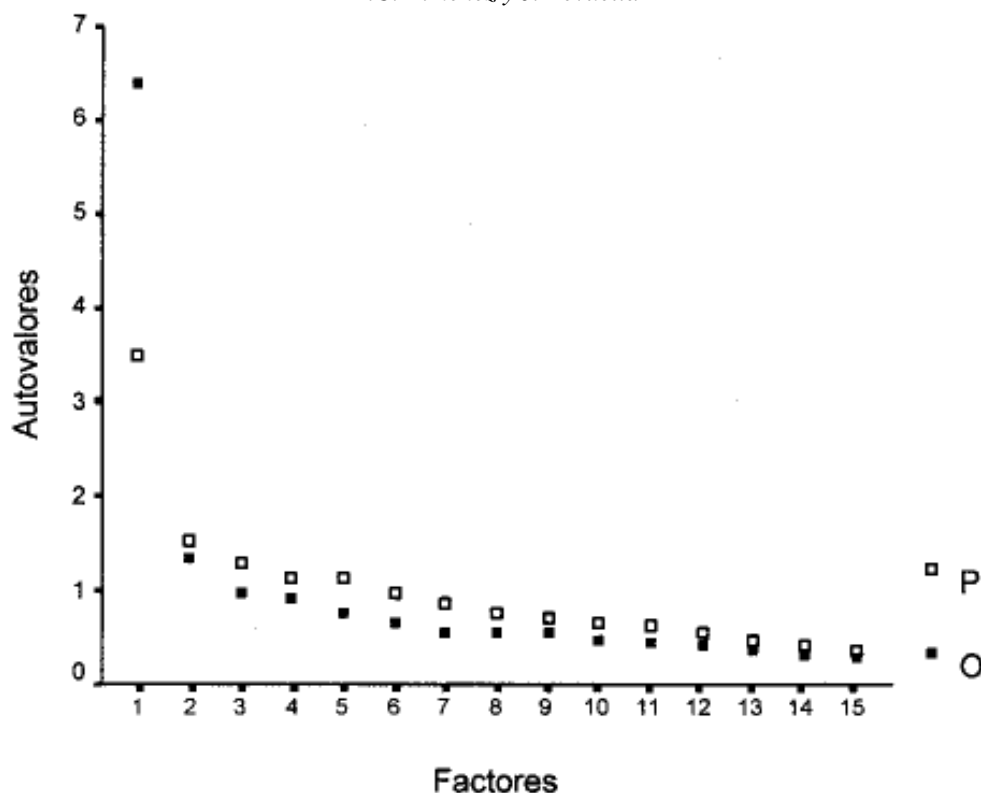


Figura 1. Representación gráfica de la comparación entre los autovalores obtenidos en los ítems de P y O.

La tabla 3 muestra el valor estimado de los parámetros del MRN para cada uno de los 15 pares de ítems gemelos, así como los residuos, calculados como la diferencia entre las proporciones observadas y esperadas de respuesta a cada alternativa. Como se indica anteriormente, los parámetros de la alternativa 1 se fijaron a cero en todos los ítems para identificar el modelo.

Puede apreciarse que los residuos son bajos para todos los ítems, lo que indica un buen ajuste del modelo. Sin embargo, aplicando el estadístico χ^2 a la diferencia entre las probabilidades marginales predichas y las observadas de cada alternativa se obtiene un valor significativo de 136.13 con 56 grados de libertad ($p < 0.01$). Este estadístico resulta muy sensible al tamaño muestral, de modo que cualquier modelo puede ser rechazado si la muestra es suficientemente grande (Marsh, Balla y McDonald, 1988), por lo que se recomienda el uso de otros índices de bondad de ajuste. Un estadístico alternativo es el cociente entre χ^2 y sus grados de libertad (gl), que toma el valor 2.43. Los criterios de aceptación para χ^2/gl oscilan entre 3 y 5 (Carmines y McIver, 1981; Wheaton, Muthen, Almin y Summers, 1977) por lo que en conjunto los resultados indican un ajuste aceptable del modelo.

Tabla 3:

Valores estimados de los parámetros y diferencia entre proporciones observadas y estimadas

ITEM	α_1	α_2	α_3	α_4	β_1	β_2	β_3	β_4	β_5	R ₁	R ₂	R ₃	R ₄	R ₅
Ítem 1	.94	2.24	3.71	4.47	1.93	2.93	2.89	1.37	.02	.01	.01	.01	.01	.01
Ítem 2	.74	1.46	2.46	3.29	1.64	2.41	2.92	2.48	.02	.02	.01	.01	.01	.00
Ítem 3	.77	1.77	3.08	3.87	1.58	2.74	2.91	1.97	.02	.01	.02	.02	.02	.02
Ítem 4	1.48	2.40	3.03	3.54	3.17	4.48	4.95	4.08	.00	.00	.00	.01	.01	.00
Ítem 5	1.66	2.82	4.13	5.20	3.37	4.76	5.61	4.79	.00	.01	-.01	.01	.01	.00
Ítem 6	2.02	3.51	5.89	7.47	3.78	5.21	5.91	4.71	-.01	.00	-.01	-.02	-.02	-.02
Ítem 7	1.14	2.81	4.26	5.29	2.54	4.59	5.33	4.56	-.03	-.01	-.02	-.01	-.01	-.02
Ítem 8	1.02	2.14	3.11	3.78	1.66	2.36	2.22	1.37	-.04	-.02	-.02	-.02	-.02	.00
Ítem 9	1.05	2.12	3.04	3.44	2.17	3.46	3.69	3.02	-.01	.00	-.03	-.03	-.03	-.03
Ítem 10	.37	1.16	1.33	.94	1.33	2.43	2.67	2.23	.01	-.01	.00	-.01	-.01	-.02
Ítem 11	.97	1.82	2.15	2.02	1.41	1.84	1.62	.67	-.01	-.02	.00	.00	.00	.00
Ítem 12	1.75	2.77	3.52	3.76	4.37	6.37	7.66	8.08	.00	-.01	-.01	.01	.01	.01
Ítem 13	1.70	3.33	5.12	6.53	2.37	3.36	3.10	2.02	.01	.01	.02	.00	.00	.02
Ítem 14	1.50	2.77	3.86	4.57	2.23	2.82	2.20	.91	.02	.01	.01	.01	.01	.00
Ítem 15	1.66	3.02	4.39	5.40	2.41	3.01	1.97	.17	.00	.00	.03	.01	.01	.01

En la tabla 3 puede apreciarse que en todos los ítems, exceptuando el 10 y el 11, se encuentra que los parámetros de sus alternativas están ordenados de modo creciente. Este resultado significa que existe un orden natural en las alternativas de respuesta en relación con el nivel de (Samejima, 1972), por lo que en estos datos podría haberse obtenido un buen ajuste con algún modelo politómico para categorías ordenadas. Esto puede hacer innecesaria la flexibilidad añadida que permite el MRN al no asumir ningún orden específico en las alternativas.

A partir de estos parámetros se estimaron los niveles de habilidad de los sujetos. La variable tiene por definición media 0 y desviación típica 1. Descompuesta en p y o se obtienen los niveles de rasgo de cada individuo en P y en O. Los estadísticos descriptivos de estas variables aparecen en la tabla 4.

Tabla 4. Estadísticos descriptivos de los niveles de rasgo.

Variable	Media	Desviación	Mínimo	Máximo
θ_p	.76	.58	-1.77	2.52
θ_o	-.78	.82	-3.35	2.39

La interpretación de este resultado es sencilla sabiendo que la distribución completa es aproximadamente la normal estandarizada. Puede verse que los sujetos perciben que sus necesidades en el trabajo son mayores que el grado en que las perciben promovidas en su organización.

La hipótesis de igualdad de medias se contrastó mediante la prueba T para muestras relacionadas, resultando un estadístico de contraste $t = 40.50$, con 590 grados de libertad ($p < 0.001$). Por tanto puede rechazarse la hipótesis de igualdad de medias en la población. El tamaño del efecto es el siguiente (Pardo y San Martín, 1998): diferencia de medias estandarizada $d = 1.66$, proporción de varianza explicada $r = 0.86$. Otro dato de interés es que la correlación entre los dos valores de rasgo de cada sujeto es $r = 0.17$ ($p < 0.001$), lo que sugiere la necesidad de utilizar un modelo más general para la estimación de parámetros, en el cual se estime la covarianza entre distintos rasgos.

En la figura 2 aparece el gráfico de dispersión de los valores de θ_p y θ_o . Cada punto del gráfico se corresponde con los dos valores estimados para un mismo sujeto. En el gráfico se aprecia como la mayoría de los sujetos tienen niveles de rasgo en P superiores a los de O, por lo que casi la totalidad de la nube de puntos cae por debajo de la bisectriz.

CONCLUSIONES

El objetivo principal de este artículo es de naturaleza metodológica: presentar un método para estimar valores de rasgo conmensurables en dimensiones diferentes. El método propuesto, basado en un modelo politómico de rasgo latente, ha resultado útil a este respecto. Los resultados obtenidos indican que dicho método proporciona niveles de rasgo distribuidos según la normal estandarizada en el conjunto de las dimensiones consideradas. Esto permite una interpretación sencilla de los niveles de cada una de las dimensiones particulares.

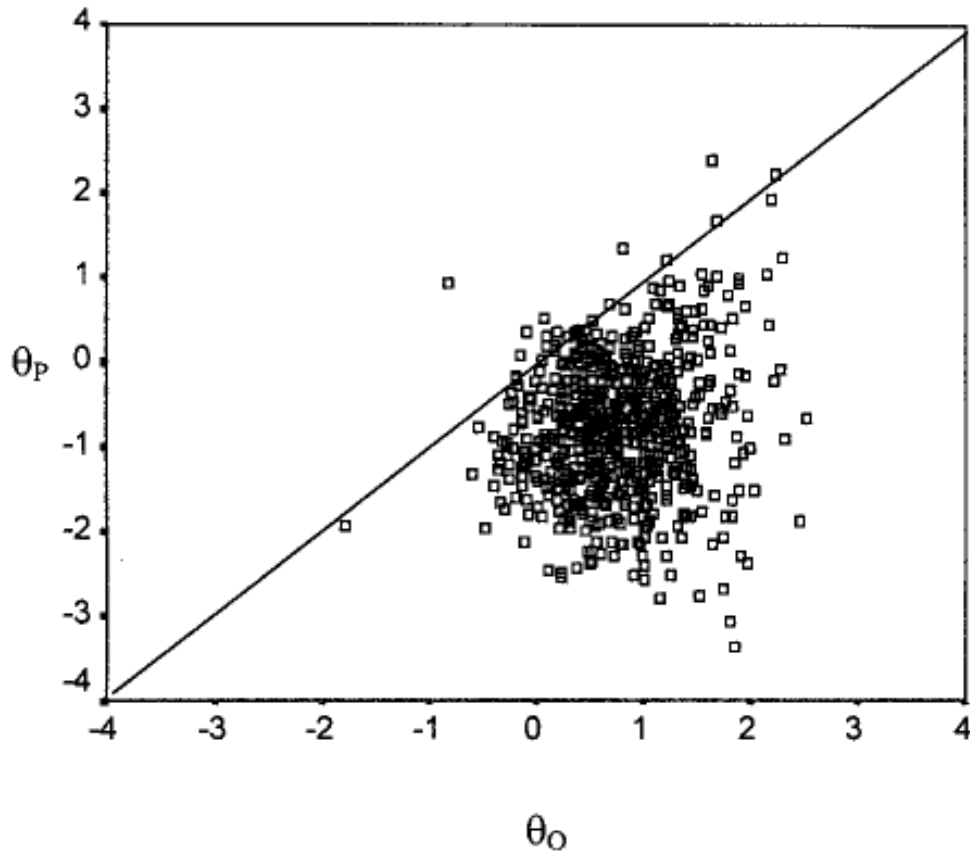


Figura 2. Gráfico de dispersión de θ_p y θ_o

El método se basa en unos supuestos fuertes que pueden limitar su utilidad práctica mientras no se obtengan versiones más generales del modelo. Estos supuestos son la independencia estocástica de y_p e y_o , la unidimensionalidad de ambas escalas, y la independencia de p y o .

Un supuesto fundamental del método propuesto es la independencia local entre las respuestas del sujeto. Esto debe entenderse como una primera aproximación al problema para lograr mayor sencillez matemática. En el cuestionario presentado resulta verosímil encontrar dos tipos de dependencia: Una entre las respuestas dadas a cada par de ítems gemelos, y otra entre los ítems de cada mitad con el resto de los ítems de dicha mitad. En fases sucesivas de la investigación debe aplicarse un modelo de TRI que recoja las dependencias entre distintos ítems para evaluar la magnitud de este efecto.

Respecto al supuesto de unidimensionalidad, puede concluirse que se cumple moderadamente en los datos analizados, aunque se han encontrado algunas diferencias en la composición de los factores y en la varianza explicada que indican la necesidad de profundizar en el análisis de la estructura factorial del cuestionario.

Una limitación del método ha sido no considerar las covarianzas poblacionales en la estimación de los niveles de rasgo. En particular, cabe esperar y así se ha encontrado, que los niveles de los distintos rasgos de cada sujeto muestren covarianzas distintas de cero, indicando que los sujetos con mayor necesidad (P) perciben también una mayor satisfacción de la misma por parte de su organización (O). En el método empleado se asume que cada nivel de rasgo procede de una normal estandarizada. Podría considerarse un modelo más general en el que el conjunto de niveles proceda de una distribución normal multivariada con matriz de covarianza no-diagonal.

El trabajo descrito debe considerarse solamente como un primer paso que todavía adolece de importantes carencias metodológicas. En particular, se comienza ajustando un modelo lineal de rasgo latente para garantizar la unidimensionalidad de un modelo no lineal, que es el que tiene verdadero interés para la interpretación de resultados. Esta forma de proceder resulta habitual en la investigación aplicada en parte por la fácil disponibilidad de software. Sin embargo, desde un punto de vista teórico sería más adecuado ajustar únicamente modelos multidimensionales de TRI para datos politómicos, y seleccionar aquel que mejor explique los resultados (Kelderman, 1996). Una investigación de esta naturaleza claramente excede los límites del presente trabajo.

El resultado más relevante del estudio empírico presentado aquí es que se ha comprobado que en esta muestra los sujetos tienen niveles de rasgo en P significativamente superiores a los de O. Esto significa que las personas sienten que sus necesidades y valores en el trabajo no están suficientemente satisfechas por la organización en que trabajan. Asimismo, el procedimiento expuesto para abordar el problema de la conmensurabilidad en las unidades de medida resulta útil para los ámbitos de la psicología en que se adoptan modelos y medidas de ajuste persona-ambiente.

ABSTRACT

Application of a politomous IRT model to obtain commensurate units measures in a person-organization fit scale. The present article explores the problem of the development of direct methods to achieve the criteria of 'commensurate units' in person-environment fit measures by applying a politomous model of latent trait. The response nominal model by Bock (1972; 1997) was used to assess the statistical goodness of fit of two P and O parallel measures of a person-organization (P-O) fit questionnaire with real data from 591 subjects. Results indicated that the proposed method is an appropriate approximation to the problem because it produces normal scores and this allows an easier interpretation of the units on P and O; However, it also presents some limitations that should be addressed in future research.

Key words: Item response theory (IRT), Politomous models of latent trait, commensurate units, person-organization (P-O) fit.

REFERENCIAS

- Carmines, E.G. y McIver, J.P. (1981). Analyzing models with unobserved variables: analysis of covariance structures. En G.W. Bornstedt y E.F. Borgatta (eds.), *Social Measurement: Current issues*. Beverly Hills: Sage.
- Carmines, E.G. y Zeller, R.A. (1979). *Reliability and validity assessment*. Londres: Sage.
- Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37, 29-51.
- Bock, R. D. (1997). The nominal categories model. En W. J. Van der Linden R.K. Hambleton (Eds.). *Handbook of Modern Item Response Theory*. New York: Springer.
- Bock, R. D. y Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters. Application of an EM algorithm. *Psychometrika*, 46, 4, 443-459.
- Chen, W. H., y Thissen, D. (1997). Local dependence indexes for item pairs using Item Response Theory. *Journal of Educational and Behavioral Statistics*, 22, 3, 265-289.
- Dempster, A. P. Laird, N. M. y Rubin D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society. Series B*, 39, 1-38.
- Edwards, J.R. (1994). The study of congruence in organizational behavior research: critique and a proposed alternative. *Organizational Behavior and Human Decision Processes*, 58, 51-100.
- Hambleton, R. K. y Swaminatham, H. (1985). *Item Response Theory: Fundamentals and applications*. Boston: Kluwer- Nijhoff.
- Hambleton, R. K. y Van der Linden, W. J. (1997). *Handbook of Modern Item Response Theory*. New York: Springer.
- Kelderman, H. (1996). Multidimensional Rasch models for partial credit scoring. *Applied Psychological Measurement*, 20, 2, 155-168.
- Lumsden, J. (1957). A factorial approach to unidimensionality. *Australian Journal of Psychology*, 9, 105-111.
- Lumsden, J. (1961). The construction of unidimensional tests. *Psychological Bulletin*, 58, 122-131.
- Masters, G. F. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 2, 149-174.
- Marsh, H.W., Balla, J.R. y McDonald, R.P. (1988). Goodness of fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103, 391-410.
- Pardo, A. y San Martín, R. (1998). *Análisis de datos en Psicología II*. Madrid: Pirámide.
- Reckase, M.D. (1979). Unifactor latent trait models applied to multifactor tests: results and implications. *Journal of Educational Statistics*, 4, 207-230.
- Rounds, J.B., Dawis, R.V. y Lofquist, L.H. (1987). Measurement of person-environment fit and prediction of satisfaction in the theory of work adjustment. *Journal of Vocational Behavior*, 31, 297-318.
- Samejima, F. (1972). A general model for free response data. *Psychometrika monograph supplement*. 37, 1.
- Thissen, D. (1991). *MULTILOG, multiple categorical item analysis and test scoring using Item Response Theory*. Chicago: Scientific Software Inc.
- Wheaton, B., Muthen, B., Almin, D. y Summers, G. (1977). Assessing reliability and stability in panel models. En D. Heise (ed.), *Sociological Methodology*. San Francisco: Jossey Bass.
- Ximénez, M.C. (1998). *El ajuste de la persona a la organización: Revisión, planteamiento de una medida y de un estudio empírico en una muestra española*. Tesis doctoral no publicada. Universidad Autónoma de Madrid.

Ximénez, M.C. y San Martín, R. (1998). El análisis de la comensurabilidad en una medida de ajuste persona-ambiente: Un estudio empírico. *Psicológica*, 19, 237-258.

(Revisión aceptada: 1/6/99)