

Evaluación de la fiabilidad del escalamiento multidimensional (EMD) a partir de matrices simétricas

Ana Delia Correa Piñero, José Miguel Díaz Gómez,
Ernesto Suárez Rodríguez y Bernardo Hernández Ruiz*

La aplicación del EMD a múltiples campos de investigación ha ido acompañada de un conjunto de estudios que intentan detectar debilidades del procedimiento y, si es el caso, sugerir vías de solución. Estos trabajos podrían dividirse en tres grupos:

Un primer grupo se centra en la resistencia del EMD a las violaciones de sus supuestos básicos: monotonicidad en el EMD no métrico y linealidad en el EMD métrico (Cohen y Jones, 1974; Green, 1975a; Spence y Ogilvie, 1973; Shepard, 1969). Otro grupo de trabajos comprueban la estabilidad de las soluciones del EMD ante diversos factores presuntamente distorsionantes: orden de presentación de los estímulos, número de estímulos, algoritmo utilizado, presencia de casos extremos, etc. (Green, 1975b; Dong, 1983; Jain y Pinson, 1976; Spence y Lewandowski, 1989). Finalmente, un tercer grupo de estudios se ha centrado en el análisis de la validez y fiabilidad del EMD: Day, Deutscher y Ryans (1976), Deutscher (1982), Humphreys (1982), Malhotra (1982, 1987), Moore y Lehmann (1982), Summers y Mackay (1976, 1977), Weksel y Ware (1967), Wilkes y Wilcox (1977). La conclusión global a extraer sobre este tema quizá no pueda ser muy clara en este momento, dadas las divergencias observadas en los resultados de unos y otros trabajos. Pensamos que se precisan más investigaciones que usen distintos estímulos, formas de obtener los datos, sujetos y ensayos de fórmulas alternativas para valorar la fiabilidad y la validez del EMD. Este es el caso de la investigación que presentamos.

Nuestro trabajo introduce dos novedades básicas con respecto a otros estudios de fiabilidad del EMD. Por un lado, el tipo de estímulos que hemos utilizado no son los habituales del EMD (colores, formas, distancias, zonas, marcas,...) sino frases relativamente cortas sobre el medio ambiente, su explotación, conservación, etc. La posibilidad de usar datos de esta índole en el escalamiento es especialmente interesante para las Ciencias Sociales, ya que con frecuencia las variables de interés no son constructos que puedan representarse con una palabra. La segunda diferencia viene determinada por el procedimiento que hemos empleado para evaluarla fiabilidad, ya que hemos adoptado un sistema basado en la comparación de pares simétricos. Los sujetos, para cada par de frases, debían valorar su similitud dos veces en una misma sesión: una vez

comparando A con B y otra comparando B con A. De esta forma se obtienen dos conjuntos de juicios equivalentes al tiempo que se evitan algunos de los inconvenientes asociados al test-retest como la mortandad experimental y efectos de variables secundarias. Se esperaba encontrar una correlación positiva y significativa entre las soluciones generadas por el EMD a partir de los dos conjuntos de juicios de similitud.

Método

Instrumentos

El instrumento de obtención de datos fue un cuestionario donde se valoraba el grado de similitud entre 21 frases relativas al medio ambiente. En el cuestionario se presenta cada frase comparada consigo misma y con las 20 restantes. El grado de semejanza de cada pareja se valora según una escala que oscila de 1 (nada parecidas) hasta 7 (iguales). Por ejemplo:

“Las soluciones simples y directas a los problemas ambientales no producen los resultados necesarios”

1	2	3	4	5	6	7
----- -----	----- -----	----- -----	----- -----	----- -----	----- -----	----- -----
nada	muy poco	poco	algo	bastante	muy	igual
parecida	parecida	parecida	parecida	parecida	parecida	

- 1 “Para solucionar los problemas ambientales es necesario cambiar nuestros hábitos de vida.”
- 2 “La creación de zonas verdes va en detrimento del crecimiento.”
- 3 “Es muy arriesgado hacer depender nuestro futuro del éxito tecnológico.”
- 4 “La satisfacción y felicidad de los ciudadanos sólo será posible si progresamos económicamente.”
- 5 ... Y así hasta presentar las 21 frases.

La comparación de todas las frases con todas las demás produce una matriz 21x21 por sujeto: la diagonal contiene las comparaciones de cada frase consigo misma (21), la mitad superior las comparaciones de cada frase con el resto (210) y la mitad inferior las mismas comparaciones invirtiendo el orden de las frases en cada pareja (210).

Sujetos

El cuestionario fue aplicado a 71 sujetos seleccionados por un procedimiento de cuotas que los distribuía equilibradamente en función del sexo y de tres grupos de edad: menos de 25, 25-50 y más de 50 años.

Procedimiento

La cumplimentación del cuestionario tuvo lugar en varias sesiones, individuales o de pequeños grupos. Las instrucciones indicaban que la tarea consistía en comparar un conjunto de

pares de frases y evaluar su grado de semejanza, insistiendo en que ésta no se refería a la estructura gramatical ni al grado de acuerdo personal con las ideas recogidas en las frases, sino al grado de similitud de dichas ideas. Los sujetos cumplieron el cuestionario en un tiempo que oscilaba de 60 a 75 minutos.

Con este procedimiento, se obtuvo para cada sujeto de la muestra una matriz cuadrada compuesta de dos mitades simétricas: [AB,BA; AC,CA...]. Cada matriz individual se descompuso posteriormente en dos matrices triangulares, obteniendo así dos subconjuntos de matrices individuales [AB; AC; AD ...] y [BA; CA; DA ...]. Que fueron escaladas independientemente mediante INDSCAL. Finalmente, se valoró la fiabilidad del ENID correlacionando las soluciones halladas en los dos subconjuntos.

Resultados

El escalamiento se llevó a cabo con 40 de los sujetos de la muestra, seleccionados al azar, debido a las limitaciones establecidas por el programa: $Nxe^2 \leq 18000$ (N: número de sujetos; e= número de estímulos) (MDS, 1981). En ambos casos (mitad inferior y superior de la matriz original) se adoptó una solución de dos dimensiones. El porcentaje de varianza total de la solución fue de 41,3 para la mitad superior y de 32,5 para la inferior. En la mitad superior, el porcentaje de varianza de la dimensión 1 fue de 78,5 y de 21,5 para la dimensión 2. En la inferior, las dimensiones 1 y 2 explicaban un 69,7% y un 30,3%, respectivamente.

Para valorar la fiabilidad de la solución se correlacionaron las coordenadas de los 21 estímulos en el espacio bidimensional obtenidas en el escalamiento de una mitad del cuestionario con las mismas coordenadas obtenidas en el escalamiento de la otra mitad. Las correlaciones producto-momento se reflejan en la tabla siguiente:

	dimensión 1 mitad superior	dimensión 2 mitad superior	dimensión 1 mitad inferior	dimensión 2 mitad inferior
dimensión 2 Mitad superior	0.40 p=0.036	1.000 *****		
Dimensión 1 Mitad inferior	0.95 p=0.000	0.44 P=0.021	1.000 *****	
Dimensión 2 Mitad inferior	0.42 p=0.027	0.84 P=0.000	0.32 P=0.075	1.000 *****

Discusión

La posición de los estímulos en cada dimensión de la mitad inferior tiene una alta y significativa correlación con su posición en la misma dimensión de la otra mitad. Los resultados

apoyan la estabilidad del AMD, en cuanto a la ubicación de los estímulos, a partir de conjuntos simétricos de juicios de similitud. Se observa que a medida que nos alejamos de la primera dimensión, disminuye la correlación intradimensión. Sería necesario profundizar en este resultado a partir de escalamientos donde se hayan obtenido más dimensiones, para comprobar si esto constituye una tendencia estable. Es probable que sea así, y ello se deba a que las sucesivas dimensiones van explicando menos porcentaje de varianza. Por otra parte, la similitud de las correlaciones obtenidas en mitades distintas entre las dimensiones 1 y 2 refuerza la idea de que ambas soluciones reflejan la misma estructura multidimensional subyacente.

La conclusión que más cabría destacar es la pertinencia del uso del EMD con estímulos más "complejos" que los usados habitualmente. No obstante, el estudio de la fiabilidad y validez del EMD precisa de más estudios donde se analice su sensibilidad al tipo de estímulos, a los instrumentos de recogida de datos y a los procedimientos usados para valorar la fiabilidad y validez. Sería preciso, paralelamente, efectuar revisiones de los resultados que se obtienen en estudios dispersos. En estas revisiones, que podrían adoptar la forma de meta-análisis, se analizarían las razones de las discrepancias de los resultados obtenidos por los distintos autores, con el fin de alcanzar conclusiones claras sobre el tema.

Referencias

- COHEN, H. & JONES, L. (1974). The effects of random error and subsampling of dimensions on recovery of configurations by nonmetric mds. *Psychometrika*, 39, 69-90.
- DAY, G., DEUTSCHER, T. & RYANS, A. (1976). Data quality, level of aggregation, and nonmetric mds solutions. *Journal of Marketing Research*, 13, 92-97.
- DEUTSCHER, T. (1982). Issues in data collection and reliability in marketing mds studies - Implications for large stimulus sets. En R.G. Golledge & J.N. RAYNER (Eds.) *Proximity and preference: problems in the multidimensional analysis of large data sets*. University of Minnesota Press, Minneapolis.
- DONG, H. (1983). Method of complete triads: An investigation of unreliability in multidimensional perception of nations. *Multivariate Behavioral Research*, 18, 85-96.
- GREEN, P. (1975A). On the robustness of mds techniques. *Journal of Marketing Research*, 12, 73-81.
- GREEN, P. (1975B). Marketing applications of mds: assessment and outlook. *Journal of Marketing*, 39, 24-31.
- HUMPHREYS, M. (1982). Data collection effects on nonmetric mds solutions. *Educational & Psychological Measurement*, 1005-1022.
- JAIN, A. & PINSON, C. (1976). The effect of order of presentation of similarity judgments on mds results: An empirical examination. *Journal of Marketing Research*, 13, 435-439.
- MALHOTRA, N. (1982). Structural reliability and stability of nonmetric conjoint analysis. *Journal of Marketing Research*, 19, 199-207.
- MALHOTRA, N. (1987). Validity and structural reliability of mds. *Journal of Marketing Research*, vol. XXIV, 164-173.
- MDS (1981). *MDS package user manual*. Program Library Unit. University of Edimburg.

- MOORE, W. & LEHMANN, D. (1982). Effects of usage and name on perceptions of new products. *Marketing Science*, 1, 351-370.
- SHEPARD, R. (1969). *Some principles and prospects for the spatial representations of behavioral science data*. Advanced Research Seminar on Scaling and Measurement, Newport Beach, California.
- SPENCE, I. & Lewandowsky, S. (1989). Robust multidimensional scaling. *Psychometrika*, 54(3), 501-513.
- SPENCE, I. & Ogilvie, J. (1973). A table of expected stress values for random rankings in nonmetrics mds. *Multivariate Behavioral Research*, 8, 511-518.
- SUMMERS, J. & MACKAY, D. (1976). On the validity and reliability of direct similarity judgments. *Journal of Marketing Research*, 13, 289-295.
- SUMMERS, J. & MACKAY, D. (1977). On establishing convergent validity: A reply to Wilkes and Wilcox. *Journal of Marketing Research*, 14, 263-265.
- WEKSEL, W. & WARE, E. (1967). The reliability and consistency of complex personality judgments. *Multivariate Behavioral Research*, 2, 537-541.
- WILKES, R. & WILCOX, J. (1977). On the validity and reliability of direct similarity judgments: A comment. *Journal of Marketing Research*, 14, 261-262.