Psicothema
2000. Vol. 12, Supl. nº 2, pp. 534-539

ISSN 0214 - 9915 CODEN PSOTEG

Análisis factorial confirmatorio de segundo orden y matrices multirrasgo-multimétodo

José M. Tomás, Amparo Oliver y Pedro M. Hontangas
Universidad de Valencia

En este trabajo se compara el modelo de análisis factorial confirmatorio de segundo orden, como alternativa a los modelos aditivos tradicionales para analizar la validez convergente y discriminante cuando se utiliza más de un indicador por combinación de rasgo-método. En concreto se comparan el modelo de rasgos correlacionados y métodos correlacionados o modelo completo y el modelo de análisis factorial confirmatorio de segundo orden. Se realizó un estudio de simulación Monte Carlo con estos dos modelos para generar los datos, utilizando una matriz multirrasgo-multimétodo con tres rasgos y tres métodos. Las variables independientes manipuladas son la correlación entre factores de método (0, .2, .4 y .6), el tamaño muestral (200 y 1000) y el modelo de análisis factorial estimado (modelo completo y modelo confirmatorio de segundo orden). En cada condición se utilizaron 100 replicaciones. Las variables dependientes consideradas son la existencia de problemas de estimación (la convergencia del proceso y la aparición de soluciones mal definidas) y el sesgo y la raíz del error cuadrático medio de las estimaciones de las saturaciones de rasgo (validez convergente) y las correlaciones entre rasgos (validez discriminante). Los resultados indican que el modelo completo es superior al modelo confirmatorio de segundo orden. Este modelo presenta muchos problemas de estimación, que lo descartan como alternativa frente a los modelos tradicionales e indican que nunca debe ser utilizado en las condiciones descritas en esta investigación.

Confirmatory factor analysis and MTMM matrices. In this paper, second-order factor analysis is compared with a first-order additive model (complete model) traditionally used to analyse convergent and discriminant validity. A simulation study was performed with the complete model and the second order model as data generators, using multitrait-multimethod matrices with three traits and three methods. Conditions of the study included several indicators per trait-method combination. Manipulated variables included sample size (200 and 1000) and correlation among method factors (0. .2, .4, .6) and the estimated confirmatory model (complete model and second order model). 100 replications were used in each condition. Dependent variables under scrutiny were: a)convergence problems, whether or not the model converged; b) ill-defined solutions; c) bias in the estimation of trait factor loadings (convergent validity) and trait correlations (discriminant validity); and d) the root mean square error in the estimation of the trait factor loadings and trait correlations. Results shown the superiority of the complete model over the second order confirmatory factor analysis. The second order model presented so many estimation problems that almost discard it as an alternative over the first order models in order to analyse multitrait-multimethod matrices included in this study.

Una de las estrategias propuestas para obtener evidencia empírica sobre la validez de constructo de un instrumento es analizar su validez convergente y discriminante utilizando un diseño en el que se miden varios rasgos con diferentes metodos —diseño MRMM— (Campbell y Fiske, 1959). La *validez convergente* se refiere al grado de acuerdo entre varias medidas del mismo constructo obtenidas por distintos métodos. La *validez discriminante* se refiere al grado de diferenciación entre distintos constructos. En términos operativos, y en palabras de los anteriores autores, existe validez discriminante cuando las relaciones entre las medidas de los diferentes constructos son bajas frente a las relaciones entre distintas medidas del mismo constructo. Junto a la validez convergente y discriminante se encuentran los *efectos de método*, que son sesgos que alteran las correlaciones entre los diferentes constructos medidos por diferentes métodos. Campbell y Fiske (1959) emplean el término método para referirse a múltiples tests del mismo constructo, múltiples ocasiones de medida, múltiples jueces, múltiples formulaciones o modificaciones de los mismos tests, etc. En un sentido más amplio, los efectos de método pueden definirse como cualquier fuente de variación sistemática en la medición al margen del constructo de interés. En consecuencia, puede hablarse simplemente de error de medida sistemático. Para señalar la importancia que tiene el error de medida sistemático, baste señalar el

Correspondencia: José M. Tomás
Facultad de Psicología
Universitat de València
46010 Valencia (Spain)
E-mail: tomasjm@uv.es

trabajo de Buckley, Cote y Comstock (1990), quienes, en el contexto de medidas de personalidad y actitud, encuentran que la varianza sistemática explica menos del cincuenta por ciento de la varianza de las medidas analizadas. El diseño MRMM puede utilizarse para analizar estos efectos de método incluso en su acepción más extensiva.

Las matrices multirrasgo-multimétodo siguen siendo muy utilizadas como diseño de investigación. El valor heurístico de las directrices presentadas por Campbell y Fiske (1959) es ampliamente reconocido, pero su aplicación operativa, el análisis de datos que se deriva del diseño, ha sido muy criticado (Schmidt y Stults, 1986; Widaman, 1985). Las críticas se centran fundamental, pero no exclusivamente, en utilizar un análisis de correlación bivariado, ya que «las diferencias entre las fiabilidades de las distintas medidas distorsionará tanto las correlaciones entre las variables como las medidas de resumen que se deriven de ellas» (Widaman, 1985, pag. 2). Los problemas asociados a la formulación de Campbell y Fiske han dado lugar a otros enfoques, como el análisis de varianza y el análisis factorial (Kenny y Kashy, 1992), siendo el más prometedor el análisis factorial confirmatorio (AFC) (Marsh y Grayson, 1995). Algunos autores afirman que solamente el AFC permite una descomposición precisa de la variabilidad debida al constructo, al error sistemático (efecto de método) y al error aleatorio (Buckley, Cote y Comstock, 1990). En el AFC existen, a su vez, distintos modelos cuya utilidad varía en función del tipo de efectos de método, el número de variables observables, el número de rasgos y métodos considerados y el tamaño de la muestra, entre un largo etcétera. Una primera clasificacion de modelos de AFC aplicados a matrices MRMM distingue entre modelos aditivos y no aditivos. Los más utilizados, hasta el momento, en la literatura aplicada han sido los aditivos, probablemente por su mayor facilidad de realización con los programas estándar de modelos de ecuaciones estructurales (EQS, LISREL). Tambien existen diferentes submodelos aditivos, pero no son más que distintas especificaciones de las relaciones entre las variables.

El primer modelo aditivo en aparecer es el modelo completo o modelo de rasgos correlacionados y métodos correlacionados (Jöreskog, 1971). Una representación del modelo completo para tres rasgos y tres métodos aparece en la figura 1(a). En este modelo cada variable observable (cuadrado en la figura) representa una medida de un constructo recogida mediante un método de medida particular (ocasión, juez…). El AFC descompone la variabilidad de cada variable observable en el efecto del constructo, el efecto del método y la unicidad. El modelo asume la independencia entre las unicidades y permite que los rasgos puedan estar correlacionados, así como los métodos, pero impone que los rasgos y los métodos sean independientes entre sí. Existen modelos particulares, entre los que merece destacar el que asume la ortogonalidad de los factores de método. Pese a sus ventajas respecto a otras técnicas, el modelo completo tiene importantes inconvenientes prácticos debido a los problemas de identificación y estimación que presenta, especialmente la falta de convergencia del proceso de estimación y la obtención de estimaciones fuera de rango (valores aberrantes como varianzas negativas o casos Heywood) (Brannick y Spector, 1990; Kenny y Kashy, 1992; Kumar y Dillon, 1992; Marsh y Bailey, 1991; Wothke, 1996).

Una alternativa al modelo completo es el modelo de unicidades correlacionadas presentado por Marsh (1988, 1989), basándose en trabajos de Kenny (1976, 1979). Los efectos de método se deducen a partir de la correlación entre las unicidades de las medidas de los rasgos que tienen en común un mismo método. Una representación gráfica de este modelo, también para tres rasgos y tres métodos, aparece en la figura 1(b). Este modelo se diferencia del modelo completo en dos aspectos fundamentales (Marsh, 1993): a) asumir que los métodos no están correlacionados (la violación de éste supuesto podría generar problemas) y b) permitir efectos de método multidimensionales, esto es, que más de un método explique una misma variable observable. El modelo de unicidades correlacionadas parece razonablemente robusto a violaciones del supuesto de independencia de los métodos (Marsh y Bailey, 1991) y, además, reduce los problemas de estimación presentes en el modelo completo (Bagozzi, 1993).

Los resultados obtenidos al comparar estos modelos aditivos apuntan hacia una superioridad o, al menos, una mayor utilidad del modelo de unicidades correlacionadas frente al modelo completo. Esta evidencia proviene de investigaciones analíticas, empíricas y de simulación, y puede considerarse sólida. No obstante, existe una limitación importante. Los resultados proceden casi exclusivamente de estudios que usan matrices multirrasgo-multimétodo «clásicas»; es decir, se trata de matrices con: un mínimo de tres

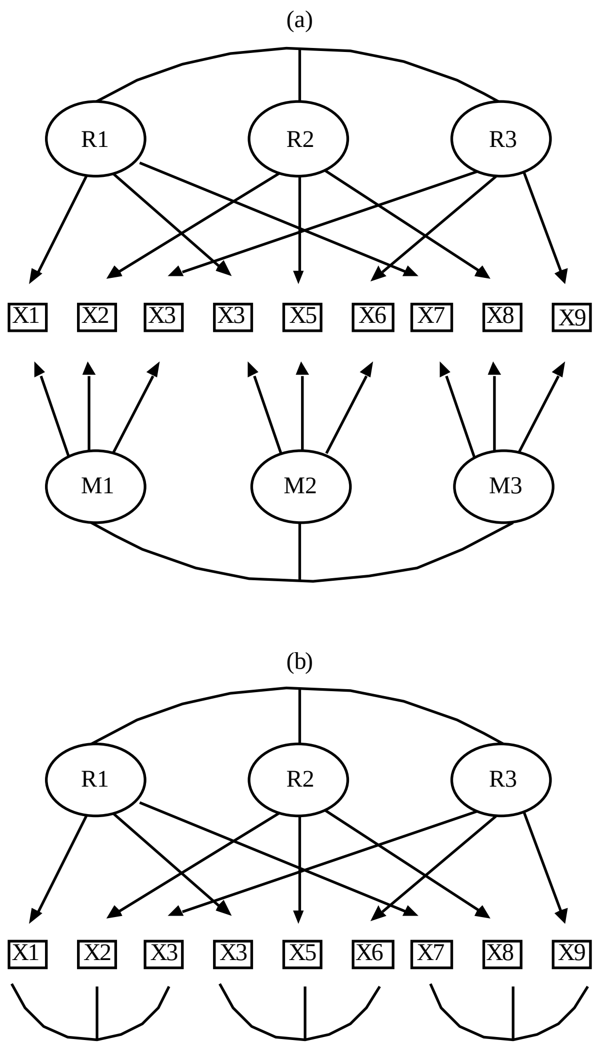

Figura 1. *Modelo completo (a) y modelo de unicidades correlacionadas (b). Notas: Las unicidades se eliminan por simplicidad; las líneas curvas representan covarianzas*

rasgos y tres métodos y un solo indicador por combinación de rasgo-método. Estas circunstancias podrían limitar de alguna manera las conclusiones obtenidas o el uso de estos modelos en condiciones aplicadas más eficientes y viables. En este sentido, Bollen y Paxton (1998) afirman que en muchas áreas aplicadas los requerimientos de las matrices multirrasgo-multimétodo clásicas pueden ser muy exigentes, no atienden adecuadamente la necesidad de estudiar los efectos de método con menor número de rasgos y/o métodos y, sobre todo, no aprovechan la información disponible en situaciones muy comunes donde hay más de un indicador por combinación rasgo-método. También Marsh (1993) plantea que no es necesario que una matriz multirrasgo-multimétodo tenga un único indicador por combinación rasgo-método; por el contrario, la presencia de mayor número de indicadores por combinación rasgo-método podría tener sus ventajas. Un procedimiento para analizar este tipo de datos es utilizar un modelo de análisis factorial confirmatorio de segundo orden (Marsh, 1993; Marsh y Hocevar, 1988). En este modelo, cuya representación gráfica puede verse en la figura 2, cada combinación rasgo-método se representa por un factor de primer orden que se infiere a partir de los indicadores múltiples de esa combinación. A partir de esa estructura de factores de primer orden, los efectos de rasgo y de método se infieren mediante factores de rasgo y de método de segundo orden de la misma forma que en el modelo completo se infieren a partir de factores de primer orden (Marsh, 1993). Esta nueva especificación no ha sido prácticamente utilizada hasta ahora y plantea la necesidad de explorar la utilidad de un modelo aditivo de este tipo para analizar matrices multirrasgo-multimétodo.

La evidencia empírica sobre el análisis de matrices multirrasgo-multimétodo con más de un indicador por combinación rasgo-método es muy limitada y proviene fundamentalmente de dos fuentes. Por un lado, se encuentran algunos estudios de validez convergente-discriminante que han utilizado el modelo completo o el modelo de unicidades (Bagozzi y Heatherton, 1994; Bollen y Paxton, 1998; Hontangas et al. 1997; Marsh, 1993; Marsh y Hocevar, 1988; Tomás y Oliver, 1999). Estos trabajos empíricos han mostrado la utilidad del modelo completo con indicadores múltiples, al estar relativamente libre de los problemas de estimación que aparecen en el caso de un solo indicador por combinación de rasgo-método. También los trabajos mencionados de Marsh señalan la superioridad teórica del modelo de segundo orden y la ausencia de problemas de estimación. Por otro lado, Tomás, Hontangas y Oliver (en prensa) comparan el modelo completo y el modelo de unicidades correlacionadas con múltiples indicadores en un estudio de simulación y obtienen que, en este contexto, el modelo completo en ciertas condiciones es superior al modelo de unicidades correlacionadas.

El objetivo del presente trabajo es analizar el comportamiento del modelo de análisis factorial confirmatorio de segundo orden como alternativa a los modelos aditivos tradicionales para analizar la validez convergente y discriminante cuando se utiliza más de un indicador por combinación de rasgo-método. Se pretende comparar las características de esta nueva modelización de matrices multirasgo-multimétodo y el modelo completo de rasgos correlacionados y métodos correlacionados, puesto que ha revelado ser mejor que el modelo de unicidades correlacionadas ante este problema de investigación en la literatura más reciente.

Método

Modelos MRMM y generación de datos

Se realizó un estudio de simulación Monte Carlo con dos modelos de referencia para generar los datos: el modelo completo y el modelo de análisis factorial de segundo orden. El tipo de matriz multirrasgo-multimétodo empleado tiene tres rasgos y tres métodos con tres indicadores por combinación rasgo-método. Los modelos se corresponden, con ligeras modificaciones en el número de variables, con los que aparecen en las figuras 1(a) y 2. Los valores de las saturaciones de rasgo fueron establecidos entre .5 y .7 (promedio=.6), las saturaciones de método entre .25 y .35 (promedio=.3), y las correlaciones entre rasgos se fijaron en .3. Los datos fueron generados con el programa EQS 5.1 (Bentler, 1995) de manera que las variables tuvieran una distribución normal multivariada.

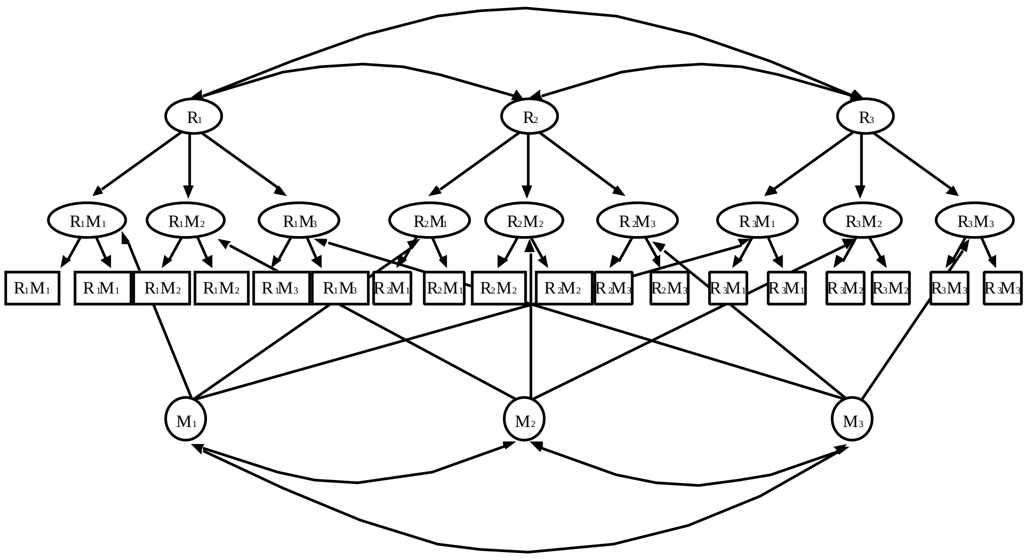

Figura 2. *Modelo factorial confirmatorio de segundo orden. Nota, las unicidades se eliminan por simplicidad; las líneas curvas representan covarianzas*

Diseño

Un diseño 4x2x(2) fue utilizado con cada uno de los modelos de generación de datos, siendo las dos primeras variables entre-sujetos (correlación entre factores de método y tamaño muestral) y la tercera de medidas repetidas (modelo de análisis factorial estimado). La magnitud de la correlación entre factores de método tiene cuatro niveles: 0, .2, .4 y .6. El tamaño de la muestra tiene dos niveles: 200 y 1000 sujetos. Y el modelo de análisis factorial estimado tiene dos niveles: modelo completo y modelo confirmatorio de segundo orden. Por tanto, tenemos un total de 32 condiciones, en cada una de cuales se utilizaron 100 replicaciones.

La estimación de los parámetros de los modelos factoriales fue obtenida con el programa EQS 5.1 (Bentler, 1995). El método de estimación seleccionado es máxima verosimilitud, dada la normalidad multivariada de los datos. El número máximo de iteraciones permitido para alcanzar la convergencia en el proceso de estimación se fijó a 100, siendo el criterio de convergencia igual a .001 y el de tolerancia igual a .000001.

Variables dependientes

Las variables dependientes consideradas pueden agruparse en tres bloques:

a) Problemas de identificación y estimación. Incluye la variable convergencia (si el modelo converge o no), y la variable soluciones bien definidas (la solución se considera bien definida cuando el proceso de estimación converge y no se presenta ninguna solución impropia).

b) Saturaciones factoriales. En el caso de las saturaciones se necesitan dos variables dependientes para la evaluación de la adecuación de la estimación, que respectivamente miden el signo y cuantía en valor absoluto de la discrepancia entre las estimaciones y sus valores poblacionales. Estas dos variables dependientes, utilizadas por Marsh y Bailey (1991), son el sesgo, medido como el promedio para cada modelo del alejamiento de la saturación estimada frente a su valor poblacional, y la precisión, operacionalizada como la raíz cuadrada de la media de las diferencias cuadráticas entre las saturaciones estimadas y su valor poblacional (RMSE, root mean square error)

c) Correlación entre rasgos. Se evalúan las mismas dos variables dependientes utilizadas en el caso de las saturaciones, sesgo y precisión, pero aplicadas a las estimaciones de las correlaciones entre los rasgos.

Resultados

Convergencia y soluciones bien definidas

El primer problema que puede presentarse en la estimación de los modelos es que la solución no converja después del número de iteraciones predeterminado. Los resultados obtenidos indican que los problemas de convergencia han sido muy poco frecuentes. El modelo completo converge en el 98.8% de los casos cuando es el modelo poblacional (modelo utilizado para generar los datos). Cuando no es el modelo poblacional, es decir, cuando es un modelo más simple pero incorrectamente especificado, converge en todas las ocasiones. Por su parte, el modelo confirmatorio de segundo orden presenta porcentajes de convergencia muy similares: converge el 98.1% de los casos, cuando no es el modelo poblacional, y el 97.5% de los casos, cuando es el modelo poblacional. Por tanto, los problemas de convergencia son prácticamente inexistentes. En todas las condiciones simuladas y con los parámetros de convergencia implicados en el análisis, las soluciones obtenidas tienden a converger en un porcentaje muy elevado en ambos modelos.

Muy diferente es la situación en cuanto a problemas de estimación. En el análisis factorial confirmatorio es posible obtener soluciones impropias (valores no razonables o imposibles) al estimar los parámetros del modelo. Un caso típico, que no el único, es la aparición de varianzas de error negativas, conocido como casos Heywood. Si en la estimación del modelo hay un problema de convergencia o una solución impropia, se considera que la solución está mal definida y el resultado no debe interpretarse. En el caso de que la solución converja y no haya soluciones impropias, la estimación del modelo resulta interpretable y se dice que la solución está bien definida. Los resultados obtenidos muestran un gran número de soluciones impropias. Así pues, aunque los problemas de convergencia son insignificantes, no ocurre lo mismo con las soluciones bien definidas. Este problema afecta mucho más al modelo confirmatorio de segundo orden que al modelo completo. El modelo completo presenta un 91.6% de soluciones bien definidas, cuando es el modelo poblacional, y un porcentaje aún mayor (100%), cuando el modelo poblacional es el modelo confirmatorio de segundo orden. Todas las soluciones mal definidas aparecen en la condición de un tamaño muestral pequeño (n= 200) y en las condiciones con factores de método oblícuos. Por otro lado, el modelo confirmatorio de segundo orden presenta graves problemas de estimación: cuando es el modelo poblacional, hay únicamente un 25% de soluciones bien definidas, mientras que cuando es el modelo incorrectamente especificado, *ninguna* solución está bien definida. Por lo tanto, los problemas de soluciones mal definidas se concentran, de forma muy acentuada, en el modelo confirmatorio de segundo orden.

El análisis más detallado de las condiciones en las que el modelo confirmatorio de segundo orden presenta problemas de estimación revela que, cuando no es el modelo poblacional, hay un 100% de soluciones mal definidas, y cuando es el modelo poblacional, existe sólo un 25% de soluciones bien definidas. Su frecuencia es muy desigual entre condiciones y niveles experimentales. En cuanto al tamaño muestral, la situación más favorable es un tamaño muestral elevado, como era de esperar. Hay un 40.75% de soluciones bien definidas con una muestra de 1000 sujetos frente al 9.25% con una muestra de 200 sujetos. También hay diferencias según la cuantía de la correlación entre métodos. El porcentaje de soluciones bien definidas es sensiblemente mayor con métodos ortogonales (45.5%, r=0) frente métodos correlacionados (17%, r=.2; 18%, r=.4; 19.5%, r=.6). Estas dos variables podrían interactuar, sin embargo el patrón de resultados de la correlación entre métodos se mantiene inalterado entre los diferentes tamaños muestrales.

Los resultados sobre las saturaciones factoriales de rasgo y las correlaciones entre rasgos se presentarán únicamente para las soluciones bien definidas del modelo completo, puesto que las soluciones mal definidas no deben interpretarse y en el modelo confirmatorio de segundo orden se han encontrado una altísima cantidad de ellas.

Saturaciones factoriales de rasgo

El sesgo y la precisión (RSME) en la estimación de las saturaciones de rasgo se presentan separadamente para el modelo completo cuando es el modelo poblacional (modelo especificado correctamente) y cuando no es el modelo poblacional (modelo especificado incorrectamente). Los resultados que describen estas dos situaciones se encuentran en la tabla 1.

Cuando el modelo completo es el modelo poblacional, la estimación de las saturaciones factoriales de rasgo es prácticamente insesgada y la precisión elevada en todas las condiciones del diseño. La ausencia de sesgo se mantiene entre los distintos tamaños muestrales, pero se aprecia una precisión algo menor cuando el tamaño muestral es pequeño (n=200). No hay diferencias aparentes entre los distintos niveles de correlación de los factores de método. Tampoco parece que los resultados sobre sesgo y precisión cambien sustancialmente en función de combinaciones particulares de las dos variables.

Cuando el modelo completo no es el modelo poblacional, no se mantiene la ausencia de sesgo, pues se produce un sesgo negativo importante igual a -.22 (el promedio de las saturaciones es .6). Por tanto, si se aplica el modelo completo, se espera una infraestimación de las saturaciones de rasgo. Estos resultados son similares, aunque con mínimas diferencias, a través de todos los niveles de las variables y combinaciones de éstos. La precisión de las estimaciones tampoco puede calificarse de adecuada, pues el error promedio a través de todas las condiciones es .27. Las diferencias entre niveles y condiciones experimentales son también poco relevantes.

Correlación entre rasgos

Los resultados relativos a las correlaciones entre rasgos se presentan en la tabla 1. En ella se describen los datos obtenidos al estimar el modelo completo cuando es el modelo poblacional que genera los datos y cuando el modelo poblacional es el modelo confirmatorio de 2º orden. En el caso de que el modelo completo sea el modelo poblacional los resultados son prácticamente iguales a los comentados sobre las saturaciones de rasgo; esto es, la estimación resulta muy precisa y prácticamente insesgada. La situación es diferente cuando el modelo completo no es el modelo poblacional. En este caso, el sesgo negativo está presente, como en el caso de las saturaciones, así como una importante falta de precisión. Los resultados a través de condiciones se mantienen constantes también, pero son sustancialmente menores en cuantía.

Conclusiones

En este trabajo hemos explorado la adecuación del modelo confirmatorio de segundo orden frente al modelo completo de rasgos correlacionados y métodos correlacionados como técnicas de análisis de matrices multirrasgo-multimétodo para estudiar la validez convergente y discriminante cuando se utiliza más de un indicador por combinación de rasgo-método. De los resultados obtenidos se desprende que el modelo completo no tiene problemas de convergencia y sólo en unos pocos casos aparecen problemas de estimación debidos a soluciones mal definidas. Por el contrario, el modelo confirmatorio de segundo orden presenta un número muy elevado de problemas de estimación, aunque tiene un nivel de convergencia razonable en la mayoría de los casos. Estos resultados permiten concluir que el modelo confirmatorio de segundo orden no es una alternativa frente a los modelos aditivos tradicionales y, por tanto, no debería ser utilizado en ninguna de las condiciones analizadas en este estudio.

Los resultados del análisis factorial de segundo orden son muy inadecuados, lo que abre la discusión sobre su porqué. La opinión de los autores a este respecto es que muy probablemente los problemas del confirmatorio de segundo orden sean debidos a la identificación del modelo y/o su complejidad. Un modelo confirmatorio de segundo orden está identificado si, por un lado, la estructura de relaciones de los factores de primer orden con los indicadores está identificada y si, por otro, la estructura de relación de los factores de segundo orden con los de primero está identificada. Si observamos nuevamente la figura 2, esto implica que los factores de primer orden presentan múltiples indicadores, situación que la investigación señala como beneficiosa para aumentar el número de soluciones convergentes y soluciones propias. En otras palabras, la parte del modelo que se corresponde con la estructura de primer or-

Tabla 1
Medias de sesgo y precisión (RMSE) de las estimaciones del modelo completo

		Modelo completo poblacional				Modelo factorial 2º orden poblacional			
		Saturaciones		Correlación entre rasgos		Saturaciones		Correlación entre rasgos	
		Sesgo	**Precisión**	**Sesgo**	**Precisión**	**Sesgo**	**Precisión**	**Sesgo**	**Precisión**
	Total	.00	.06	.00	.06	-.22	.27	-.11	.13
Tamaño muestral	200	-.01	.08	.01	.09	-.23	.28	-.12	.14
	1000	.00	.03	.00	.04	-.22	.26	-.10	.11
Correlación factores método	0	.00	.05	.00	.05	-.23	.28	-.13	.14
	.2	.00	.06	-.01	.06	-.23	.28	-.13	.14
	.4	-.01	.06	.00	.07	-.22	.27	-.11	.12
	.6	.00	.06	.00	.07	-.21	.26	-.08	.10
Tamaño muestral y correlación factores método	200, .0	.00	.07	.00	.07	-.23	.29	-.14	.16
	200, .2	-.01	.08	-.01	.08	-.23	.29	-.13	.16
	200, .4	-.01	.08	.00	.09	-.22	.28	-.12	.14
	200, .6	.00	.08	.03	.10	-.22	.27	-.09	.12
	1000, .0	.00	.03	.00	.03	-.22	.27	-.12	.12
	1000, .2	.00	.03	.00	.04	-.23	.27	-.12	.13
	1000, .4	.00	.03	-.01	.04	-.22	.26	-.10	.11
	1000, .6	.00	.04	-.01	.05	-.21	.25	-.07	.08

den efectivamente se beneficia en términos de estimación de los indicadores múltiples. No así la estructura de segundo orden; aquí los factores de segundo orden presentan un solo indicador por combinación rasgo-método, y por lo tanto todos los problemas asociados en el modelo completo a la presencia de un indicador único. Estos problemas son muy serios, y pueden hacer que el modelo sea muy difícilmente identificable y/o estimable, con el problema añadido de que los posibles problemas de estimación de la estructura de primer orden se *añaden* a los problemas de estimación de la estructura de segundo. El resultado global puede calificarse, como se ha visto, de calamitoso. El trabajo de Marsh (1993) no encontraba para una única matriz empírica problemas en el factorial de segundo orden, pero tampoco encontraba problemas en el modelo completo y, además, la matriz presentaba cuatro rasgos y cuatro métodos, lo que aumentaba la ratio de indicadores a factores, otro aspecto beneficioso para encontrar soluciones bien definidas.

Los resultados del modelo completo son, por el contrario, muy adecuadas. No presenta prácticamente sesgo alguno, la precisión es adecuada y los problemas de estimación aparecen con muy baja frecuencia. Evidentemente estos resultados se mantienen cuando el modelo completo es el poblacional, y no cuando es un modelo incorrectamente especificado, donde los problemas de estimación continúan sin ser un problema, pero el sesgo es elevado y la precisión de la estimación baja. No obstante, cabe señalar que el correspondiente modelo poblacional (el de segundo orden) presentó mayor cuantía de sesgo y una mayor falta de precisión. En conclusión, el modelo completo ha mostrado como en Tomás et al. (en prensa) que es un modelo muy adecuado para su aplicación en los casos en que existan indicadores múltiples de cada combinación rasgo-método.

Los resultados de cualquier estudio de simulación se circunscriben a las condiciones simuladas, a las condiciones bajo estudio, por lo que todas las conclusiones son específicas. No obstante señalan en direcciones ya apuntadas en la literatura para otras condiciones, y que quedan, por lo tanto, por explorar. En concreto, sería necesario extender las condiciones de simulación a matrices multirrasgo-multimétodo con diferentes números de rasgos y métodos, como por ejemplo cuatro rasgos y cuatro métodos, como los utilizados por Marsh (1993) para estudiar la generalización de los resultados aquí apuntados.

Agradecimientos

Este trabajo se enmarca en el proyecto PB 98-1483 del Programa Sectorial de Promoción General del Conocimiento subvencionado por la DGES, y en sendos proyectos concedidos a los dos primeros autores bajo convocatoria SEUID (MEC)-Royal Society de Londres 98.

Referencias

Bagozzi, R. P. (1993). Assessing construct validity in personality research: Applications to measures of self-esteem. *Journal of Research in Personality*, 27, 49-87.

Bagozzi, R. P., & Heatherton, T. F. (1994). A general approach to representing multifaceted personality constructs: Application to state self-esteem. *Structural Equation Modeling, 1,* 35-67.

Bentler, P. M. (1995). *EQS structural equations program manual.* Encino, C. A.: Multivariate Software, Inc.

Bollen, K. A. & Paxton, P. (1998). Detection and determinants of bias in subjective measures. *American Sociological Review, 63,* 465-478.

Brannick, M. T., & Spector, P. E. (1990). Estimation problems in the block-diagonal model of the multitrait-multimethod matrix. *Applied Psychological Measurement, 14,* 325-339.

Buckley, M- R., Cote, J. A. & Comstock, S. M. (1990). Measurement errors in the behavioral sciences: The case of personality/attitude research. *Educational and Psychological Measurement, 50,* 447-474.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Hontangas, P. M., González-Romá, V., Lloret, S., & Ferreres, D. (1997). *A study of the method effects on the factorial validity of the General Health Questionnaire (GHQ-12).* 10th European Meeting of the Psychometric Society, Santiago de Compostela.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika, 52,* 99-111.

Kenny, D. A. (1976). An empirical application of confirmatory factor analysis to the multitrait-multimethod matrix. *Journal of Experimental Social Psychology, 12,* 247-252.

Kenny, D. A. (1979). *Correlation and causality.* New York, Wiley.

Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin, 112,* 165-172.

Kumar, A., & Dillon, W. R. (1992). An integrative look at the use of additive and multiplicative covariance structure models in the analysis of multitrait-multimethod data. *Journal of Marketing Research, 29,* 51-64.

Marsh, H. W. & Grayson, D. (1995). Latent variable models of multitrait-multimethod data. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, issues, and applications.* California, Sage.

Marsh, H. W. & Hocevar, D. (1988). A new, more powerful approach to multitrait-multimethod analyses: An application of second-order confirmatory factor analysis. *Journal of Applied Psychology, 73,* 107-117.

Marsh, H. W. (1988). Multitrait-multimethod analyses. In J. P. Keeves (Ed.), *Educational research methodology, measurement and evaluation: An international handbook.* Oxford, Pergamon Press.

Marsh, H. W. (1989). Confirmatory factor analysis of multitrait-multimethod data: Many problems and a few solutions. *Applied Psychological Measurement, 13,* 335-361.

Marsh, H. W. (1993). Multitrait-multimethod analyses: Inferring each trait-method combination with multiple indicators. *Applied Measurement in Education, 6,* 49-81.

Marsh, H. W., & Bailey, M. (1991). Confirmatory factor analysis of multitrait-multimethod data: A comparison of alternative models. *Applied Psychological Measurement, 15,* 47-70.

Schmitt, N., & Stults, D. N. (1986). Methodology review: Analysis of multitrait-multimethod matrices. *Applied Psychological Measurement, 10,* 1-22.

Tomás, J. M. & Oliver, A. (1999). Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling, 6,* 84-98.

Tomás, J.M., Hontangas, P.M. & Oliver, A. (en prensa). Linear confirmatory factors models to evaluate multitraint-multimethod matrices: the effects of number of indicator and correlation among methods. Multivariate Behavioral Research.

Whothke, W. (1996). Models for multitrait-multimethod matrix analysis. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques.* New Jersey, LEA.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement, 9,* 1-26.