chapter 3

Analysis of multitrait multimethod matrices

R. KÖLTRINGER

A review of the psychological literature by Schmitt and Stults (1986) revealed that the following procedures are often used to analyze multitrait multimethod (MTMM) matrices:
- exploratory factor analysis
- smallest space analysis (SSA)
- analysis of variance, including
  a. nonparametric alternatives
  b. confirmatory factor analysis (CFA)

Of all these analytical strategies, only the CFA approach leads to numerical estimates of validity, method, and residual variance components of the indicators. Therefore the other procedures will not be described in detail.

THE GENERAL CFA MODEL FOR MTMM-DATA

Consider n persons who are rated using m different methods of measurement on t traits. This research design results in a (mt x 1) vector of observed variables **x**, which can be seen as a linear function of t + m factors ξ as follows:

$$\mathbf{x} = \Lambda \xi + \delta \qquad (1)$$

where
- Λ is a (mt x (t + m)) matrix of factor loadings
- ξ is a ((t + m) x 1) vector of t trait and m method factors
- δ is a (mt x 1) vector of residuals.

Assuming that the residuals are uncorrelated with each other and with the factors, equation (1) gives the standard factor analytic decomposition

$$\Sigma = \Lambda \Phi \Lambda' + \Theta_\delta \tag{2}$$

where
- Σ is an (mt xmt) observed covariance (or correlation) matrix
- Λ is defined as in (1)
- Φ is a (t + m) x (t + m) covariance (or correlation) matrix of the t + m factors
- Θ_δ is an (mt x mt) diagonal matrix of residual variances.

In order to simplify further discussion it is useful to partition Φ as

$$\Phi = \begin{bmatrix} \Phi_{tt} & \Phi_{tm} \\ \Phi_{mt} & \Phi_{mm} \end{bmatrix}$$

where
- Φ_{tt} is a (t x t) symmetric submatrix of Φ that contains trait factor variances and covariance
- Φ_{mm} is an (m x m) symmetric submatrix of Φ that contains method factor variances and covariances
- Φ_{mt} (= Φ'_{tm}) is an (m x t) rectangular submatrix of Φ that contains covariances of the m method factors with the t trait factors

and Λ as

$$\Lambda = [\Lambda_t \mid \Lambda_m]$$

where
- Λ_t is an (mt x t) submatrix of Λ that contains zeros, except for the loadings of observed variables on their respective trait factors
- Λ_t is an (mt x m) submatrix of Λ that contains zeros, except for the loadings of the observed variables on their respective method factors.

Assuming that the factor analytic model of Σ represented by equation (2) is valid and that scores on the mt observed measures follow a multivariate normal distribution, Jöreskog (1969, 1971) developed procedures for obtaining maximum likelihood estimates of all model parameters in Λ, Φ, and Θ_δ. The significance of each of the parameters in the model may be tested by forming a z-value of the parameter estimate by dividing its asymptotic standard error. In addition to the test of each individual parameter estimate, maximum likelihood estimation yields an overall χ^2 goodness-of-fit test, which is a test of the difference in fit between a given model and a completely saturated model that perfectly reproduces the data.

Problems of the CFA-approach

Besides the difficulties inherent in covariance structure analysis as a whole (e.g. the rather restrictive distributional assumptions), CFA models for MTMM data have several additional problems. The full CFA model for MTMM data, as represented in equation (2), is only identified for $t \geq 3$ and $m \geq 3$. However, as Kenny (1976) pointed out, the model can, with further restrictive assumptions, be used to test MTMM matrices with smaller dimensions. It can be shown that the use of other methods applicable to smaller matrices requires the same (restrictive) assumptions (Alwin, 1974).

Even under the condition $t \geq 3$ and $m \geq 3$, trait-method covariances are empirically underidentified. Applications in which parameters in Φ_{mt} were estimated resulted in

- factor loadings or factor intercorrelations greater then 1.00 for standardized data,
- the presence of Heywood Cases (estimates of unique variance close to zero), and
- very large standard errors for some parameters (because of high intercorrelations of the parameter estimates).

Constraining Φ_{mt} to be a null matrix leads to orthogonality of the trait and method factor space, a condition that is highly desirable since it results in a decomposition of the variance of each indicator into an additive combination of trait, method, and residual (random error) variance. Our present knowledge about the empirical nature or practical importance of trait-method interactions is very limited. It is possible that small trait-method correlations do not effect the trait and method factor loading estimates substantively but exaggerate the residual variances. In order to provide guide-lines for the interpretation of CFA models for MTMM-data, simulation studies on this point are required.

When $t \geq 3$ and $m \geq 3$ and all the parameters in Φ_{mt} are fixed to zero, the CFA model for MTMM data may still be empirically underidentified because of overfactoring (Rindskopf, 1984). It is well known that at least three measures of each factor are needed to identify orthogonal factor models. Therefore, in the case of a CFA-model with a $t = 3$ and $m = 3$ MTMM-design, Heywood cases and related problems are very likely if some factor intercorrelations and factor loadings of the same factors are close to zero.

To ensure nonnegativity of unique variances in Θ_δ a procedure outlined by Rindskopf (1983) can be used: The technique requires the observed measures to be treated as y-variables and the specification of exogenous latent variables (ξ), whose variance is fixed to one, as trait, method, and residual variables. Instead of the usual parameterization of MTMM models, the factor loadings and then the residual effects are estimated in Γ, while the matrices Θ_δ, Θ_ε, and Φ are not used. Since it is not the unique variances that are directly estimated but the effects of the unique factors, the residual variances, which are the squares of these effects, will always be positive. However, it is clear, that an empirically underidentified model with some large standard errors and highly correlated parameter estimates will still be underidentified if the described parameterization is used.

Some researchers have suggested the inclusion of a general factor in the model represented in equation (2). This general factor can - in the presence of both trait and method factors - account for dimensions which inflate all the correlations in a given MTMM matrix (e.g. a general trait dimension or a response set factor). But the interpretation of such a general factor is rather indeterminate. It therefore seems to be preferable not to allow for this specification.

Representing the trends present in MTMM data is a complicated undertaking with many subjective components. Several authors who analyzed the same published MTMM matrix arrived at directly contradictory conclusions (Widaman, 1985).

HIERARCHICALLY NESTED MODELS FOR MTMM DATA

Most of the problems mentioned above can be circumvented by the strategy of nested model testing. Specifically, the probability of accepting a wrong model with a good fit and the danger of overfactoring may be minimized by the systematic comparison of various alternative models. In order to reduce the intuitive dimension of the stepwise testing procedure, Widaman (1985) proposed an array of models for MTMM matrices, based on three possible structures of trait and method factors:
- no trait (or method) factors;
- trait (or method) factors with fixed intercorrelations: either 0, indicating a high level of discriminant validity, or 1, indicating total lack of discriminant validity, and
- trait (or method) factors with freely estimated intercorrelations.

Cross-classifying of the two structures results in the taxonomy of models for MTMM data presented in table 1.

table 1:
Taxonomy of structural models for MTMM-data
(following Widaman 1985)

TRAIT STRUCTURE	METHOD STRUCTURE			
	no method factors	m method factors (fixed unit correlations)	m method factors (fixed zero intercorrelations)	m method factors (freely estimated intercorrelations)
no trait factors	**H0**: null model	1 general method (not clearly interpretable)	m methods only (orthogonal)	**A**: m methods only (oblique)
t trait factors (fixed unit inter-correlations)	1 general trait (not clearly interpretable)	2 general factors (not identified)	1 general trait + m methods (orthogonal)	**B**: 1 general trait + m methods (oblique)
t trait factors (fixed zero inter-correlations)	t traits only (orthogonal)	t traits (orthogonal) + 1 general method	t traits (orthogonal) + m methods (orthogonal)	t traits (orthogonal) + m methods (oblique)
t trait factors (freely estimated intercorrelations)	**H1**: t traits only (oblique)	t traits (oblique) + 1 general method	**H2**: t traits (oblique) + m methods (orthogonal)	**H3**: full model t traits (oblique) + m methods (oblique)

Note that table 1 is demarcated into nine areas. A particular model is nested within all other models separated by at least one line of demarcation lying to the right and/or below the model in question.

Fitting and comparing all of the models listed in table 1 might lead to the selection of a best-fitting model. But following this strategy is rather cumbersome. Widaman (1985) preferred another rationale, which is based on parsimony. With this approach it is reasonable to start with a highly restricted (but theoretically possible) model and relax restrictions only if a model fails to represent adequately a set of data:

H_0 Null Model:
Hypothesis: lack of correlation among observed variables.
Parameters: θ_δ (Λ_t, Λ_m, and ϕ_{mt} fixed to zero; $\phi_{tt} = \mathbf{I}$, $\phi_{mm} = \mathbf{I}$).

H_1 First alternative Model:
Hypothesis: covariation among observed variables is due only to trait factors and their intercorrelations.
Parameters: Λ_t,ϕ_{tt}, and θ_δ; (Λ_m, and ϕ_{mt} fixed to zero, $\phi_{mm} = \mathbf{I}$).

H_2 Second alternative Model:
Hypothesis: (orthogonal) method factors are also present).
Parameters: Λ_t,ϕ_{tt}, Λ_m, and θ_δ (ϕ_{mt} fixed to zero, $\phi_{mm} = \mathbf{I}$).

H_3 Full Model:
Hypothesis: (oblique) trait factors and (oblique) method factors account for the observed correlations.
Parameters: Λ_t,ϕ_{tt}, Λ_m, ϕ_{mm}, and θ_δ (ϕ_{mt} fixed to zero).

There is at least one problem concerning the above test strategy: the importance of trait factors for the reproduction of the MTMM matrix is tested against a highly restrictive Null Model (H_0). It can be argued that the probability of accepting the alternative model (significant trait factors) is high, even if the alternative model is wrong.

Nested model comparisons

In order to test the degree of convergent validity (and, if necessary the degree of discriminant validity) models A and B in table 1 must also be estimated. Assume that the Full Model H_3 represents the data quite well on statistical and practical grounds. Then, if the fit of H_3 (t oblique trait factors and m oblique method factors) is substantially better than the fit of model A (no trait factors, m oblique method factors), there is a significant amount of covariance among measures uniquely explained by the trait factors - this represents *convergent validity.*

To test for *discriminant validity*, the fit of model H_3 should be compared with the fit of Model B, which has fixed perfect intercorrelations among trait factors (and m oblique method factors). Comparison of the fit of the models H1 (oblique trait factors only) and H_2 (oblique trait and orthogonal method factors) gives a test statistic for the significance of the *method variance* among measures. Finally, the *discriminability of method effects* can be tested by comparing the models H_2 (t oblique traits and m orthogonal method factors) and H_3 (t oblique traits and m oblique method factors).

Table 2 summarizes the described nested model comparisons for MTMM matrices. As can be seen from this table, the coefficient Q (Jöreskog, 1974) is preferred as the a test statistic for the model comparisons (other possible measures are described in Loehlin, 1987, pp. 69-71). Q is defined as the ratio of decline in χ^2 to change in degrees of freedom; $Q > 2.0$ indicates that the improvement of fit is substantial and not merely the result of chance:

$$Q = \frac{\chi^2(H_x) - \chi^2(H_y)}{df(H_x) - df(H_y)}$$

Since subsequent models in the test series do not always differ by the same degrees of freedom, the PRE-coefficient Delta, proposed by Bentler and Bonett (1980) should not be applied. But this measure is very useful to assess the level of practical fit of a given model in addition to its statistical fit.

$$\text{Delta} = \frac{\chi^2(H_o) - \chi^2(H_x)}{\chi^2(H_o)}$$

The results of the model comparisons summarized in table 2 are "ideal" in the following sense: first, the proportion of covariation among observed measures uniquely representable as convergent validity (e.g. due to trait factors) is significant. Second, there is a pronounced degree of discriminant validity. Third, the covariation explained by method factors is also quite high.

Anomalies in the proposed test sequence most often occur when unnecessary factors are included in the model. As Rindskopf (1984) pointed out, overfactoring is present when a factor has

1. no large loadings, or

table 2: Nested model comparisons for MTMM data

MODELS

M_0	specifications	χ^2	df [1]	P	anomalies [2]
H_0	no trait factors no method factors	....	36	....	....
H_1	t traits only (oblique)	....	24	....	....
H_2	t (oblique) traits, m (orthogonal) methods	....	15	....	....
H_3	t (oblique) traits, m (oblique) methods	....	12	....	....
A	m methods only (oblique)	....	24	....	....
B	1 general trait, m (oblique) methods	....	15	....	....

MODEL COMPARISONS (predicted outcomes for "normal" MTMM matrices)

comparison		testing for:	Q	accepted model [3]
H_0	H_1	significance of trait factors	....	H_1
H_1	H_2	significance of method factors	....	H_2
H_2	H_3	discriminability of method effects	....	H_3
A	H_3	convergent validity	....	H_3
B	H_3	discriminant validity	....	H_3

[1] For t=3 and m=3.

[2] Space for indicating the occurrence of Heywood Cases etc.

[3] $Q > 2$

2. only one large loading, or
3. only two large loadings, and close to zero correlations with all other so-called real factors (factors with two or more large loadings)

In all of these situations, the model will be empirically underidentified:

- In the first case, the factor can have correlations with other factors over a fairly wide range without substantially changing the reproduced covariance matrix.
- In the second case, the nonzero loading cannot be separated reliable from residual variance, and the result can be an impossibly large loading accompanied by a negative unique variance.

- In the third case, the lack of correlation between that factor and other factors leads to underidentification. (A one factor model with two indicators is not identified!).

It should be noted that "large loading" and "statistically significant loading" are not necessarily the same in this context. A factor which causes anomalies in the test sequence may have three or four statistically significant, but numerically low loadings!

Example

The application of the proposed test series for the analysis of MTMM matrices will now be demonstrated with data from a pilot study for the international research project Quality of Attitude Questions conducted in Austria, 1987. A sample of 179 persons from Vienna were asked about their "job satisfaction" (t_1), "satisfaction with their financial situation" (t_2), "satisfaction with their sexual situation" (t_3), "satisfaction with their social contacts" (t_4), and their "overall happiness" (t_5). The responses of each par-

table 3: Correlations among 15 answers of 179 respondents to questions concerning t=5 traits, each measured with m=3 methods

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
m=1																
t_1	x_1	1														
t_2	x_2	.463	1													
t_3	x_3	.239	.274	1												
t_4	x_4	.383	.285	.347	1											
t_5	x_5	.445	.447	.536	.436	1										
m=2																
t_1	x_6	.826	.426	.233	.307	.402	1									
t_2	x_7	.456	.720	.141	.233	.372	.548	1								
t_3	x_8	.168	.125	.786	.263	.439	.275	.180	1							
t_4	x_9	.231	.175	.253	.556	.326	.314	.247	.370	1						
t_5	x_{10}	.396	.375	.513	.371	.665	.382	.399	.483	.359	1					
m=3																
t_1	x_{11}	.751	.388	.178	.359	.405	.821	.458	.193	.379	.430	1				
t_2	x_{12}	.420	.668	.205	.301	.422	.513	.721	.142	.234	.385	.553	1			
t_3	x_{13}	.257	.226	.726	.267	.490	.332	.211	.743	.341	.482	.362	.362	1		
t_4	x_{14}	.325	.260	.277	.650	.365	.385	.261	.340	.629	.384	.529	.453	.466	1	
t_5	x_{15}	.302	.305	.261	.202	.486	.312	.318	.271	.425	.590	.424	.448	.440	.390	1

ticipant to three (identical) questions were measured on each of these five dimensions. The three questions differed only with respect to the scales used: a five-point scale (m_1), a nine-point scale (m_2), and a magnitude scale (m_3).

The intercorrelations of the 15 indicators resulting from this research design are presented in table 3. All analyses were performed using the LISREL VI program (Jöreskog and Sörbom, 1983).

The results of the nested model comparisons are given in table 4. The outcomes presented in this table can be interpreted as follows:

1. Both trait and method factors uniquely explain a highly significant proportion of variance.

table 4: Nested model comparisons for the matrix in table 3

MODELS

M_0	specifications	χ^2	df	P	anomalies
H_0	no trait factors no method factors	1997.73	105	.000	-
H_1	t=5 traits only (oblique)	312.63	80	.000	-
H_2	t=5 (oblique) traits, m=3 (orthogonal) methods	107.12	65	.001	-
H_3	t=5 (oblique) traits, m (oblique) methods	157.44	65	.000	[1]
A	m=3 methods only (oblique)	1402.01	90	.000	-
B	1 general trait, m=3 (oblique) methods	915.95	75	.000	TD(10) < 0 [2]

MODEL COMPARISONS

comparison		testing for:	Q	accepted model
H_0	H_1	significance of trait factors	67.4	H_1
H_1	H_2	significance of method factors	13.7	H_2
H_2	H_3	discriminability of method effects	-	H_2
A	H_3	convergent validity	51.8	H_2
B	H_3	discriminant validity	80.8	H_2

[1] For model H_3 (t oblique traits, m oblique methods), iterations do not converge. Therefore, the method factor intercorrelations were fixed to 1 in order to test for the discriminability of method effects.

[2] Parameter not significant

2. The best fitting model H_2 with 5 oblique traits and 3 orthogonal methods (though statistically rejectable) is acceptable on practical grounds: Delta = .946.
3. The parameter estimates of H_2 cannot be fully trusted because the method factor intercorrelations are empirically underidentified. The reason for this problem lies in (mostly) significant but numerically low loadings in the nine-point scale method factor (m_2). The corresponding five loadings and their standard errors were estimated to be .156 (.063), .212 (.079), .274 (.086), .202 (.096), and -.097 (.111). So, the dilemma is that this factor cannot be deleted without substantially increasing the goodness-of-fit statistic, whereas including this factor leads to biased estimates of valid, method, and residual variance components, if the method factors are correlated.

CONCLUSIONS

There is no obvious simple solution for this dilemma, except the inclusion of additional real factors, which can make the iterative estimation procedure more stable. In the original analysis of the given MTMM-matrix (see Költringer and Kluscarits, 1988), five factors were added: four (perfectly measured) socio-demographic variables and one response set dimension (extreme score tendency). Within the resulting (good fitting) model, slightly higher method effects and rather high method factor intercorrelations were estimated: $r(m_1, m_2) = .549$ (.156), $r(m_1, m_3) = .498$ (.131), and $r(m_2, m_3) = .651$ (.105).

In general, the more traits and methods that are included in the MTMM-design, the lower the probability of less than three "large" loadings per factor, i.e., the greater the chance of detecting low or moderate method (or trait) variance components when actually only low or moderate method (or trait) effects are present.

REFERENCES

Alwin, D. (1974). An analytic comparison of four approaches to the interpretation of relationships in the multitrait-multimethod matrix. In H.L. Costner (Ed.), *Sociological methodology 1974.* San Francisco: Jossey-Bass.

Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 87-104.

Jöreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrica, 34,* 183-202.

Jöreskog, K.G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrica, 36,* 109-133.

Jöreskog, K.G. (1974). Analyzing psychological data by structural analysis of covariance matrices. In C. Atkinson et al. (Eds.), *Contemporary developments in mathematical psychology.* San Francisco: Freeman.

Jöreskog, K.G., & Sörbom, D. (1983). *LISREL VI users guide.* University of Uppsala: Department of Statistics.

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

Költringer, R., & Kluscarits, W. (1988). *Subjektive Sozialindikatoren und Messfehlerproblematik: Ein kritischer Beitrag zur gegenwärtigen Praxis der Lebensqualitätsforschung.* Wien (unpublished).

Loehlin, J.C. (1987). *Latent variable models.* Hilldale, London: LEA.

Rindskopf, D. (1983). Parameterizing inequality constraints on unique variances in linear structural models. *Psychometrica, 48,* 73-83.

Rindskopf, D. (1984). Structural equation models: empirical identification, Heywood cases, and related problems. *Sociological Methods and Research, 13,* 109-119.

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

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