Open access • Journal Article • DOI:10.1111/J.2044-8317.1992.TB00975.X

A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model — Source link

Bengt Muthén, David E. Kaplan

Institutions: University of California, Los Angeles, University of Delaware

Published on: 01 May 1992 - British Journal of Mathematical and Statistical Psychology (John Wiley & Sons, Ltd)

Topics: Sample size determination and Standard error

Title

A Comparison of Some Methodologies for the Factor Analysis of Non-Normal Likert Variables: A Note on the Size of the Model

Authors
Bengt O. Muthén
David Kaplan

Publication Date
2011-10-24

British Journal of Mathematical and Statistical Psychology (1985), **38**, 171–189 *Printed in Great Britain*

A comparison of some methodologies for the factor analysis of non-normal Likert variables

Bengt Muthén and **David Kaplan**

This paper considers the problem of applying factor analysis to non-normal categorical variables. A Monte Carlo study is conducted where five prototypical cases of non-normal variables are generated. Two normal theory estimators, ML and GLS, are compared to Browne's (1982) ADF estimator. A categorical variable methodology (CVM) estimator of Muthén (1984) is also considered for the most severely skewed case. Results show that ML and GLS chi-square tests are quite robust but obtain too large values for variables that are severely skewed and kurtotic. ADF, however, performs well. Parameter estimate bias appears non-existent for all estimators. Results also show that ML and GLS estimated standard errors are biased downward. For ADF no such standard error bias was found. The CVM estimator appears to work well when applied to severely skewed variables that had been dichotomized. ML and GLS results for a kurtosis only case showed no distortion of chi-square or parameter estimates and only a slight downward bias in estimated standard errors. The results are compared to those of other related studies.

1. Introduction

In practice, factor analysis is often carried out on variables which are highly skewed and/or kurtotic and frequently are not observed on a continuous, interval scale. This paper addresses the issue of factor analysis in cases of ordered, five-category Likert scales. The paper considers treating them as interval scale normal and interval scale non-normal variables. It also discusses the use of methodology specifically developed for categorical variables in such instances. Issues of variable discreteness, skewness, and kurtosis will be studied. In Section 2 a limited statistical framework is given. Within this, Section 3 describes the design and purpose of a small Monte Carlo study and Section 4 relates the present research to previous work. The results are given in Section 5. Section 6 gives a concluding discussion.

2. Statistical framework

Assume a factor analysis model for p continuous response variables y^*,

$$\mathbf{y}^* = \boldsymbol{\nu} + \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon}, \tag{1}$$

which with ordinary specifications gives rise to the covariance structure

$$\boldsymbol{\Sigma}(\mathbf{y}^*) = \boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}' + \boldsymbol{\Theta}. \tag{2}$$

Let $\mathbf{y}$ be a p-dimensional vector of observed response variables. Usually, we specify $\mathbf{y} = \mathbf{y}^*$, but we will here emphasize cases where $\mathbf{y}$ contains ordered categorical

#12

variables, so that for variable $i (i = 1, 2, \ldots, p)$

$$y_i = \begin{cases} C_i - 1, & \text{if } \tau_{i,C_i-1} < y_i^* \\ C_i - 2, & \text{if } \tau_{i,C_i-2} < y_i^* \leqslant \tau_{i,C_i-1} \\ \vdots & \\ 1, & \text{if } \tau_{i,1} < y_i^* \leqslant \tau_{i,2} \\ 0, & \text{if } y_i^* \leqslant \tau_{i,1} \end{cases} \quad (3)$$

where the τs are threshold parameters and C_i is the number of categories for variable i. The multinomial distribution for the categorical y variables can be deduced by integration over $\mathbf{y}^*$, where in this paper we limit ourselves to the assumption of multivariate normality for the vector $\mathbf{y}^*$. With the specification of (1), (2) and (3), we will say that the factor analysis model (its covariance structure, rather) 'holds for y^*'.

The above modelling may be relevant for the situation of ordinal variables, for example, as encountered with five-category Likert scales. We will be particularly concerned with variables that depart from the zero skew and kurtosis of normal variables. While each y^* may be normal, the histograms for each y may for instance be quite skewed, e.g., due to inappropriate question wording. As an example, consider a five-category Likert variable with observed frequencies for the different y categories of 5, 5, 5, 10 and 75 per cent. With an underlying normal y^*, this means that the y variable is 'censored' in the sense that the right-most y category lumps together all the different subjects having $y^* > \tau_4$; the observed y variation is censored relative to the true underlying variation of y^*.

With five-category Likert scales, the y variables are often scored by consecutive integers, say $0, 1, 2, 3, 4$. The analysis then proceeds as if these y variables are interval scaled, assuming that $y = y^*$ with the covariance structure of (2). Strictly speaking, the model cannot be correct unless there are infinitely many categories for each y, since the right-hand side (RHS) of (1) can give rise to infinitely many values for each dimension. If we accept the approximation of discrete ys with continuous y^*s (RHS of (1)), two cases are of interest. First, given that a certain covariance structure holds true for y^*, the same structure will in general not hold for the ys; see, for instance, Olsson (1979), also discussed below. Second, a certain covariance structure may hold for the ys and not for multivariate normal y^*s. If the latter is the case, and the ys are skewed, this can be interpreted by (1) where the RHS produces skewed variables for instance with residuals that are normal, but factors that are skewed. In this second case, the covariance structure (2) does not hold for normal y^*s, and we will instead say that the model holds for y.

With y variables such as in the above Likert scale example with extreme skews it should further be noted that the ordinary measures of association, covariances and Pearson product moment correlations, may not be suitable. Not only are the ys in this example discrete, but they are also limited in range with strong censoring. In such cases, a linear model for ys may be unrealistic, and a non-linear model, such as the one using the y^*-formulation, may be more appropriate.

In some special cases, the covariance structure of (2) may hold for both normal y^*s and non-normal ys, the latter generated by (3) and again scored $0, 1, 2, \ldots$. Consider for example a single-factor model that holds for multinormal y^*s. Assume that all y^* have the same variances and loadings so that all correlations are equal and the thresholds are the same across variables, i.e. the ys have the same distributions. For all pairs of ys we then have the common correlation $\rho_y = a\rho_{y^*}$, where a will be called the attenuation factor. Consider determining the metric of the factor by fixing a loading to a non-zero value. Then the covariance matrix for the ys can be expressed as

$$\Sigma(\mathbf{y}) = \boldsymbol{\lambda}_y \psi_y \boldsymbol{\lambda}_y' + \boldsymbol{\Theta}_y, \tag{4}$$

where

$$\begin{aligned} \boldsymbol{\lambda}_y &= \boldsymbol{\lambda}, \\ \psi_y &= \psi \sigma_y^2 a, \\ \theta_y &= \sigma_y^2 (1 - \lambda^2 \psi a). \end{aligned} \tag{5}$$

Here, σ_y^2 is the common variance for the ys and θ_y is the common diagonal element of $\boldsymbol{\Theta}_y$. In this case, $\Sigma(\mathbf{y}^*)$ and $\Sigma(\mathbf{y})$ have the same covariance structure, but with different parameter values. We will then say that the model holds for both y and y^*.

In the simulation study to be reported below, data are generated according to the special single-factor model just described (see Section 3). Here, we are in a position to study how well the $\Sigma(\mathbf{y}^*)$ structure can be recovered. Although the normal y^* model is the one that actually generated the data, such data may at the same time be used to illustrate how various estimators perform in the estimation of $\Sigma(\mathbf{y})$ structures. We are then considering effects of non-normal ys on estimation. The y^* model is only used as a means of conveniently producing non-normal ys, but our interest is in recovering $\Sigma(\mathbf{y})$. A consistent estimator would consistently estimate the parameters of this covariance structure, so that an infinitely large sample will give a perfect model fit. Finite sample biases may however occur. Although both structures hold in the above special case, this can also be viewed as illustrating cases where no simple underlying normal factor structure exists for y^*s. Whether or not such an underlying y^* structure may exist for any given data set is a matter of statistical testing and will not be considered here (see, e.g., Muthén, 1984).

In passing, consider again the special single-factor model above for normal y^*s. With a choice of different threshold values for different variables, the distributions would vary across the ys. It is well known (see, e.g., Olsson, 1979) that despite equality of y^* correlations, this would yield different correlations across pairs of y variables; there would be a differential correlation attenuation. In general, this means that although the factor analysis structure $\Sigma(\mathbf{y}^*)$ holds true, $\Sigma(\mathbf{y})$ would not obey the same structure even with different parameter values. Here, an infinitely large sample would not give a perfect fit for the factor analysis structure hypothesized; the structure is distorted by differential attenuation. In this case, a consistent estimator applied to the covariance matrix of y would not only be affected by non-normality but also by the model structure being incorrect.

3. Design and purpose of the study

3.1. *Estimators*

In this paper we will consider four different estimators. The first one is the traditional maximum-likelihood (ML) estimator, which assumes continuous, multivariate normal observed variables; see, for example, Jöreskog (1969, 1977). Here the fitting function is

$$F_{\mathrm{ML}} = \log|\Sigma| - \log|\mathrm{S}| + \mathrm{tr}(\mathrm{S}\Sigma^{-1}) - p, \tag{6}$$

where S is the sample covariance matrix. Both Σ and S refer to the distribution of the *y*s in our study. The second one is the generalized least-squares (GLS) estimator, which assumes continuous variables for which all fourth-order cumulants are zero; see, for example, Jöreskog & Goldberger (1972), Browne (1974). Here the fitting function is

$$F_{\text{GLS}} = \operatorname{tr}(\mathbf{I}-\mathbf{S}^{-1}\boldsymbol{\Sigma})^2. \quad (7)$$

The requirement of zero fourth-order cumulants is practically the same as that of multivariate normality. Hence, we will refer to both ML and GLS as normal theory estimators. Under multivariate normality, ML and GLS have the same asymptotic properties (Browne, 1974).

The third estimator is an interesting and less explored one that has recently been proposed for continuous non-normal variables, not requiring zero fourth-order cumulants. This is the so-called asymptotically distribution-free (ADF) generalized least-squares estimator of Browne (1982). The fitting function for both GLS and ADF (which is also a GLS estimator) can be written

$$F_{\text{WLS}} = (\mathbf{s}-\boldsymbol{\sigma})'\mathbf{W}^{-1}(\mathbf{s}-\boldsymbol{\sigma}), \quad (8)$$

where the WLS subscript stands for GLS or ADF, s is the $p(p+1)/2$ vector of distinct elements of S, $\boldsymbol{\sigma}$ contains the corresponding elements of Σ, and $\mathbf{W}$ is a consistent estimator of the asymptotic covariance matrix of s. Browne (1982) pointed out that the general form for covariances between covariances in a sample of size N is

$$(N-1)\operatorname{cov}(s_{ij}, s_{kl}) = \sigma_{ik}\sigma_{jl}+\sigma_{il}\sigma_{jk}+\frac{N-1}{N}\kappa_{ijkl}, \quad (9)$$

where κ_{ijkl} is a fourth-order cumulant (see, e.g., Kendall & Stuart, 1977, pp. 340, 342) Equation (9) gives the weight matrix for Browne's ADF estimator. Since in GLS zero fourth-order cumulants are assumed, the weight matrix becomes a function of the covariance matrix only, with considerable computational savings.

The fourth estimator explicitly takes into account the categorical nature of the observed variables to be considered here, using the model specification of (1), (2) and (3) with multivariate normal y^*s. This is the generalized least-squares estimator of Muthén (1978, 1984), which will be referred to as the CVM (categorical variable methodology) estimator. Using limited information from all pairs of *y*s, CVM avoids the use of Pearson product moment correlations and instead fits the model to the estimated latent correlations of the y^*s. The fitting function can be written as in (8), where s and $\boldsymbol{\sigma}$ would refer to y^* correlations and with a specific choice of weight matrix.

3.2. *Simulation design*

We will limit the study to ordered, five-category variables and the special single-factor model described in Section 2, studying cases where a factor model is true for both $\Sigma(\mathbf{y}^*)$ and $\Sigma(\mathbf{y})$. In all cases we will use a simple four-variable, single-factor model for $\Sigma(\mathbf{y}^*)$, where all loadings are 0·7, the factor variance is 1, and the residual variances are all 0·51. Hence, the y^*-variances are all 1 and the y^* correlations are all equal and medium-sized, 0·49. We will generate multivariate normal y^* vectors according to $\Sigma(\mathbf{y}^*)$, using GGNSM (IMSL, 1982), categorizing y^*s at certain threshold values to form *y*s. We will limit ourselves to large-sample properties and will hence choose a sample size of $N = 1000$ throughout. The sampling procedure will be repeated 25 times. In the analyses we choose to determine the metric of the factor by fixing the first loading to 0·7. Fixing a loading as opposed to the factor variance is a common approach of confirmatory factor analysis and structural equation modelling.

Let us consider measures of non-normality for the y variables. In the univariate case, (9) becomes

$$(N-1)\,\mathrm{var}\,(s_{ii}) = 2\sigma_{ii}^2 + \frac{N-1}{N}\kappa_{iiii}. \tag{10}$$

Note that

$$\gamma_{2,i} = \frac{\kappa_{iiii}}{\sigma_{ii}^2}, \tag{11}$$

using $\gamma_{2,i}$ to denote the univariate (excess) kurtosis coefficient for variable i, standardized to be zero for a normal variable; $\gamma_2 = \beta_2 - 3$ (cf. Kendall & Stuart, 1977, p. 88). From a univariate perspective, equations (10) and (11) indicate that it is the kurtosis which gives information on the importance of the error of the GLS assumption when variables are non-normal. Let skewness be denoted γ_1, where $\gamma_1^2 = \beta_1$. Note that large skewness implies large kurtosis, since (cf. Kendall & Stuart, 1977, pp. 88, 95) $\gamma_2 > \gamma_1^2 - 2$ must hold. In our study we will consider not only univariate skewness (γ_1) and kurtosis (γ_2), but also the multivariate counterparts suggested by Mardia (1970); see also Mardia (1974), Mardia & Zemroch (1975), Browne (1982). In the sample, Mardia's measures are defined as

$$b_{1,p} = N^{-2} \sum_{i=1}^{N} \sum_{j=1}^{N} d_{ij}^3, \tag{12}$$

$$b_{2,p} = N^{-1} \sum_{i=1}^{N} d_{ii}^2, \tag{13}$$

with

$$d_{ij} = (\mathbf{y}_i - \bar{\mathbf{y}})'\mathbf{S}^{-1}(\mathbf{y}_j - \bar{\mathbf{y}}), \tag{14}$$

where $\bar{\mathbf{y}}$ is the sample mean vector. Denoting the corresponding population measures $\beta_{1,p}, \beta_{2,p}$, we have $\beta_{1,1} = \beta_1$ and $\beta_{2,1} = \beta_2$. A symmetric distribution has $\beta_{1,p} = 0$ and a multivariate normal distribution has $\beta_{2,p} = p(p+2)$. In our study, we will consider multivariate skewness defined as $\beta_{1,p}$, which reduces to γ_1^2 for the univariate case (reported as (skewness)^2 in Table 1 below), and, as recommended by Browne (1984), multivariate relative kurtosis defined as $\beta_{2,p}/p(p+2)$, which has the value 1 for multivariate normal distributions, and reduces to $(\gamma_2+3)/3$ for the univariate case.

The data generation procedure will be carried out for each of five cases, according to different choices of thresholds for the variables. In our opinion, these choices represent Likert scale data commonly encountered in the social and behavioural sciences; see Table 1. Case 1 corresponds to a desirable situation where all y variables are symmetric with univariate skewness, kurtosis, and fourth-order cumulants close to those of the normal distribution. Here we can study the effects of categorization without interference of skewness or kurtosis. In cases 2, 3 and 4, increasing degrees of skewness and kurtosis are introduced. Case 4 represents a situation where the y variables are strongly 'censored', i.e. we observe a 'piling up' of observations at one of the extreme categories. All of these cases are presumably commonly encountered in real data. More severe skews than case 4 seem to be relatively rare (see, however,

Table 1. Descriptive statistics of y for all cases

	Case 1		Case 2		Case 3		Case 4		Case 5	
	T[b]	C[b]	T[b]	C[b]	T[b]	C[b]	T[b]	C[b]	T[b]	C[b]
		5		5		3		5		5
	−1·645		−1·645		−1·881		−1·645		−1·645	
		21		10		6		5		7·5
	−0·643		−1·036		−1·341		−1·282		−1·150	
		48		20		13		5		75
	0·643		−0·385		−0·772		−1·036		1·150	
		21		30		30		10		7·5
	1·645		0·385		0·050		−0·674		1·645	
		5		35		48		75		5
Variance	0·820		1·360		1·100		1·248		0·550	
Skewness	0·000		−0·742		−1·217		−2·028		0·000	
$(\text{Skewness})^2$	0·000		0·551		1·615		4·113		0·000	
Kurtosis	0·004		−0·334		0·846		2·898		2·785	
Univariate relative kurtosis	1·001		0·889		1·282		1·966		1·928	
Fourth-order cumulant	0·003		−0·618		1·024		4·510		0·843	
Multivariate skewness[a]	0·082		1·796		5·625		15·421		0·131	
Multivariate relative kurtosis[a]	0·989		1·026		1·277		1·892		1·580	
Correlation	0·437		0·436		0·416		0·345		0·352	
Attenuation factor	0·892		0·890		0·850		0·703		0·718	

[a] A sample estimate based on a random sample of 1000 (four variables).
[b] T = Threshold; C = Category.

Section 4.2.1 below), and this is also deemed to be the case for situations with large skews of opposite sign (given a positive correlation). While we recognize the importance of kurtosis deviations from zero as discussed above, in cases 2, 3 and 4, high kurtosis values may be seen as arising incidentally as an effect of skewness; highly correlated skew and kurtosis values would seem to frequently be the case in practice. However, a zero skew, kurtosis only case, case 5, is also included for comparison. A kurtosis close to that of case 4 was attempted, representing a leptokurtic case with a very high percentage of responses in the neutral category. While the multivariate kurtosis value for case 5 is not as high as for case 4, we think that case 5 already borders on what can be viewed as realistic for Likert scales in the social and behavioural sciences.

3.3. *Purpose of the study*

It should be noted that ML, GLS and ADF are all consistent estimators of parameters of $\Sigma(\mathbf{y})$. Consistency holds true even if observed variables are not multivariate normal, as long as $\Sigma(\mathbf{y})$ holds true (Browne, 1974). On the other hand, CVM consistently estimates parameters of $\Sigma(\mathbf{y}^*)$, although only if the y^*s are multivariate normal. Given that the respective prerequisites are fulfilled, each estimator produces asymptotically normal estimates, a large sample chi-square test of model fit, and large sample standard errors of estimates. The aim of this paper is to study to what extent these quantities behave appropriately in the case of ordered, categorical variables. Hence, in line with the usage of confirmatory (as opposed to exploratory) factor analysis and structural equation modelling, we place a heavy emphasis on chi-square and standard errors. Also, as is common here, the sample covariance rather than correlation matrix will be analysed for ML, GLS and ADF.

Our first objective is to study how well the normal theory estimators ML and GLS perform regarding chi-square and variability measures and in the estimation of $\Sigma(\mathbf{y})$ parameters. In particular, how skewed and kurtotic can the *y*s be for these estimators to still behave approximately correctly? A second objective is to study how well the more optimal ADF estimator manages the analysis of $\Sigma(\mathbf{y})$ parameters for non-normal *y*s that are ordered categorical. A third objective is to study how ML, GLS and ADF perform in the estimation of $\Sigma(\mathbf{y}^*)$ parameters. The CVM estimator will only be utilized in one situation, case 4. Here, we have for simplicity chosen to work with dichotomized *y*s, the highest category versus all others, since not much information would seem to be lost by collapsing categories. The $\Sigma(\mathbf{y}^*)$ structure is correct for CVM and the fourth objective will be to study its sampling properties.

Admittedly, this is a very limited Monte Carlo study, both with respect to different models studied, different sample sizes, and number of replications. This is due to cost restrictions, particularly due to relatively heavy computations for ADF and CVM. Nevertheless, the study should give a good deal of important information. Judging from related research to be discussed below, certain generalizations beyond our five cases seem possible. For example, the assumption of exactly equal univariate *y* distributions is unrealistic, but may serve as a 'prototype' for realistic data (cf. Boomsma, 1983). Before presenting the results, we will now give a brief overview of some related research and point out what our study contributes.

4. Overview of some related research

4.1. *ML results*

4.1.1. *Fuller & Hemmerle (1966); Boomsma (1983).* An early study, which is only marginally related to ours, is that of Fuller & Hemmerle (1966). Considering certain two-factor models, the authors were concerned with the effects of non-normal continuous data in the estimation via normal theory ML. Problems of ordered categorical variables were not addressed. Various non-normal distributions (uniform, truncated normal, *t* distribution, triangular and bimodal) were chosen for each of the two factors and the residual. The authors concluded that both with respect to chi-square and estimates the ML procedure is 'relatively insensitive to departure from normality'. Arriving at this conclusion, the authors used five variable, one degree of freedom models, and a sample size of $N = 200$. The design must be regarded as rather weak, however, since only a single replication was used for each case. Also, in most cases, the deviations from normality in the *y*s were not very large (see Fuller & Hemmerle, 1966, p. 258 and also Olsson, 1979).

Recently, Boomsma (1983) performed an interesting and very rigorous Monte Carlo study to investigate, among other things, effects of non-normality on ML estimation. Boomsma pointed out the distinction, discussed above, between a certain covariance structure holding true for $\Sigma(\mathbf{y}^*)$ or for $\Sigma(\mathbf{y})$. He decided on the investigation of confirmatory factor analysis and structural equation models holding true for $\Sigma(\mathbf{y})$ in order to study effects of non-normality only, without the confounding effect of categorization of y^*s into *y*s. As in our case, Boomsma chose to study non-normality in the form of ordered categorical *y*s for which a simple structure holds. (These were in fact obtained by a categorization of normal y^*s, although without a simple structure holding for these y^*s; see Boomsma, p. 147). In line with Olsson, Boomsma was concerned with the effects of skewness. Using a sample size of $N = 400$ and 300 replications of each case, he generated data according to four models with number of variables ranging from 6 to 10 and the size of correlations varying across the full range. For each such model he generated data according to various combinations of number of categories and skewness. A symmetric case, such as our case 1, was always included, and the skewed cases exhibited mixed skews ranging up to the value four. Throughout these cases, Boomsma in fact found very little bias in parameter estimates. However, for the more skewed cases, particularly those with most correlations medium-sized or high, he found that the empirically determined variation of the estimates was higher than that estimated by the standard errors of the estimates. Also, for those same cases, he found the chi-square measure of model fit to reject the true model much too often. Effects of number of categories and categorization with no skewness seemed to be very minor. He concluded:

> On the basis of our findings we shall not dissuade researchers to apply maximum likelihood estimation in structural equation modelling, when the observed variables are discrete but symmetric. However, we do not recommend to use such a procedure when the median (or mean) absolute value of the skewnesses of the observed variables is larger than 1·0 (approximately), because it would affect crucial elements of statistical estimation (confidence intervals for parameters, correlations among parameter estimates), and model fitting.

Again, we should note that this conclusion only pertains to the ML estimation of Σ(y structures holding true. In our study, the confounding of non-normality and categorization that Boomsma was concerned with is avoided, since a certain factor model is true for both Σ(y) and Σ(y*). We can therefore add to Boomsma's results by considering results in relation to both models at the same time, as explained in Section 2. Also, our study includes GLS, ADF and CVM, in addition to ML and studies case 5 with zero skew.

4.1.2. *Olsson (1979)*. An important study is that of Olsson (1979). Olsson was particularly interested in the effects on factor analysis of categorizing multinormal y^*s into ordered categorical ys, scored as $0, 1, 2, \ldots$. Further theoretical developments have been made by Mooijaart (1983), who related the y correlations to the underlying y^* parameters in a simplified way, and McDonald (1974), discussing similar issues in the context of so-called difficulty factors with dichotomous ys. Olsson chose to study a single-factor model for the y^*s, six and twelve variables, and a variety of number of categories and threshold combinations with an emphasis on strongly skewed variables and variables of opposite skew. Olson studied the effects of distortions in the Σ(y) structure as deviating from the true Σ(y*) structure. He did not use Monte Carlo simulated data, but instead derived the population Σ(y)s for analysis by ML. This corresponds to results for infinitely large samples. The population approach enabled Olsson to cover a large variety of situations, from which he concluded that 'classification, as we have defined it, may give rise to a substantial lack of fit'. Also, 'the skewnesses of the variables, rather than the number of scale steps, seems to be a major determinant of lack of fit of the factor model'. Important biases were obtained for factor loadings. The number of variables seemed to have little effect whereas increasing size of true loadings, and thereby correlations, made things worse. Olsson attributed the discrepancy regarding robustness as compared to the Fuller & Hemmerle study to his use of more severely non-normal variables. We should note here that for situations where all y-variables have the same distribution, Olsson obtains a perfect fit to the $\Sigma(\mathbf{y})$ structure (the variances of the y^*s were standardized to one). Hence, lack of fit, as measured in the chi-square metric, only reflects distortions in the $\Sigma(\mathbf{y})$ structure relative to the true $\Sigma(\mathbf{y}^*)$ model. This is noteworthy, since with simulated data we are able to also study lack of fit for models where Olsson would obtain perfect fit. We furthermore add to Olsson's work on $\Sigma(\mathbf{y}^*)$ estimation by including GLS, ADF and CVM in addition to ML, and studying effects of categorization in the zero skew case, case 5.

4.2. *GLS and ADF results*

4.2.1. *Some case studies.* In the above reported studies, ML was the only estimator considered. In the literature of today, however, there are also some limited results on the performance of the GLS and ADF estimators. Jöreskog & Goldberger (1972) noted that the GLS estimator produced lower error variances than ML in a certain factor analysis of a real data example with nine variables, three factors, and $N = 200$. Browne (1982) noticed that 'ADF estimators of parameters in covariance structures tend to be noticeably biased below the true values'. He compared ML and ADF estimation using a data set (Huba *et al.*, 1981) with 13 five-category variables, three factors, and $N = 1634$. This data set is further analysed by ADF in Huba & Bentler (1983). For these variables skewness values ranged up to around ten, with an average of about four, while kurtosis values ranged up to 127 with an average of about 27. The multivariate relative kurtosis value was 4·19. We may view this data set as an extreme version of our case 4. The ADF chi-square was less than a third of the ML chi-square. No large differences were found among estimated loadings, while error variances consistently obtained lower estimated values by ADF than by ML. Furthermore, estimated standard errors were generally smaller for ML than for ADF. Bentler (1983) also found the ML chi-square value to be considerably higher than the ADF value for a simulated data set with zero skewness and positive kurtosis values. Huba & Harlow (1983) and Huba & Tanaka (1983) compared ML, GLS and ADF in applications with non-normal data. Huba & Harlow (1984) also used the CVM estimator for comparison. The general outcome of these real data comparisons of estimators was a large degree of similarity.

4.2.2. *Tanaka (1984); Browne (1984).* Tanaka (1984) performed a small Monte Carlo study which addressed the performance of ML and ADF but not GLS for variables with high positive kurtosis. Using a two-factor model with three continuous variables loading on each factor, yielding an 8 degree of freedom model, he generated data with sample sizes of $N = 100, 500, 1500$, using 20 replications. Tanaka generated his data as a mixture of two multivariate normals with equal and zero means but covariance matrices differing by scaling with a diagonal matrix with different sized diagonal elements. He obtained univariate kurtosis values ranging up to about 7, with a mean value of about 5. The skewness values were however zero. Hence, his data are rather different from those considered by Olsson, Boomsma and also the data of Huba *et al.* (1981), considered by Browne (1982) and Huba & Bentler (1983). Tanaka's correlations were small to medium-sized. Excluding the $N = 100$ case, he found that ML chi-square values were considerably overestimated, whereas ADF chi-square behaved appropriately. The loadings and error variances were however considerably underestimated by both ML and ADF. This was more serious for error variances than for loadings. Interestingly, ADF was markedly more biased than ML. Tanaka also concluded that the ML estimated standard errors were markedly biased downwards, whereas this bias was much less pronounced for ADF. It should be note that Tanaka's results on parameter estimate bias for ML are in contrast with the results of Boomsma. By contrasting the high skew/high kurtosis case 4 with the zero skew/high kurtosis case 5, our study adds to Tanaka's. Note, however, for reasons mentioned in Section 3.2, that the case 5 kurtosis value is much lower than the typical value of Tanaka's. Also, he did not study ordered, categorical variables. We also add to Tanaka's kurtosis study by including the GLS estimator for case 5.

Browne (1984) included a small Monte Carlo study which among other things considered the properties of the ML and ADF estimators. Browne studied data from a rescaled multivariate chi-square distribution, such that each variable had univariate population skewness and kurtosis of 2 and 6, respectively. The multivariate relative kurtosis coefficient was 1·93. Here, 20 replications were used, with sample size 500. Two models were considered: an eight variable intraclass correlation model with 34 degrees of freedom and an eight variable, single factor model, parameterized as a correlation structure, with 20 degrees of freedom. All population correlations were 0·5. While the ADF chi-square test performed well, ML gave much too high values. Regarding parameter estimates, ML showed no bias for either model, while ADF exhibited a clear bias for the intraclass model. There was a clear downward bias in the estimated standard errors for ML, and a similar but less noticeable trend was observed for ADF. We note that Browne's results are in line with Tanaka's, except for Tanaka's finding of ML parameter bias. We may view Browne's data as a more severely non-normal version of our case 4, although he did not consider ordered, categorical variables. We add to Browne's study by including GLS and analysing ordered, categorical variables.

5. Results

As a first step, a quality check was made of the data generation process. In Tables 2 and 3 are given results for the analysis of 25 replications with $N = 1000$ for multivariate normal y^*s, that is, not introducing the categorization to ys. Results ar only presented for GLS and for Case 1 and Case 4 y^*s. The degrees of freedom of th model is two. In terms of chi-squares, Table 2 shows that GLS for case 1 tends to slightly overestimate chi-square, i.e. the true y^* model is rejected slightly too often, while for case 4, chi-square is slightly underestimated. This gives an indication of th degree of imprecision in the generating process. The crudeness of the data-generatio process should be kept in mind when considering chi-squares for the cases studied. From Table 2 it appears that parameter estimates are behaving as expected; i.e. no important, consistent bias can be detected. Furthermore, a comparison of expected standard errrors using the population $\Sigma(\mathbf{y}^*)$ with mean estimated standing errors of estimates gives another quality control on the data-generating process. For samples of this size no substantial differences are expected, and none is found. However, the crudeness of a Monte Carlo study with only 25 replications shows in the empirical standard deviations of the estimates. These are often somewhat smaller than the other two measures of variability. To remedy this, one may have to use up to, say, 300 replications in all situations, which would be too costly. This crudeness must be borne in mind when considering the results to be presented on variability of estimates.

To partially address the concern about the small number of replications we studi cases 1–5 each with 300 replications and $N = 1000$, where due to cost consideration we only applied the normal theory GLS estimator to each sample covariance matri: Relative to 25 replications the results showed very little difference in general trend:

Table 2. Parameter estimates and their variability: GLS on y^* variables

		Case 1 y^*		Case 4 y^*	
Variable	True value	Estimate[a]	Variability[b]	Estimate	Variability
			Loadings		
1	0·700	0·700 (—)	— — —	0·700 (—)	— — —
2	0·700	0·698 (−0·3)	0·039 0·039 0·032	0·704 (0·6)	0·039 0·039 0·035
3	0·700	0·691 (−1·3)	0·039 0·038 0·033	0·688 (−1·7)	0·039 0·039 0·036
4	0·700	0·685 (−2·1)	0·039 0·038 0·035	0·695 (−0·7)	0·039 0·039 0·031
			Error variances		
1	0·510	0·498 (−2·4)	0·031 0·031 0·022	0·516 (1·2)	0·031 0·031 0·028
2	0·510	0·503 (−1·4)	0·031 0·031 0·034	0·511 (0·2)	0·031 0·031 0·021
3	0·510	0·508 (−0·4)	0·031 0·030 0·033	0·510 (0·0)	0·031 0·031 0·029
4	0·510	0·520 (2·0)	0·031 0·031 0·025	0·517 (1·4)	0·031 0·031 0·032
			Factor variance		
	1·000	1·032 (3·2)	0·089 0·090 0·064	1·027 (2·7)	0·089 0·091 0·077
			Chi-square		
Mean		2·793		1·603	
Variance		6·157		1·557	
Reject freq.[c]		2		0	
			Skewness/kurtosis		
Multivariate skewness[d]		0·100		0·113	
Multivariate relative kurtosis[d]		0·973		1·007	

[a] In parentheses is given percentage under- or over-estimation of the true value.
[b] The three entries under Variability are:
—expected (true) standard error of estimate;
—mean of estimated standard errors;
—empirical standard deviation of estimates.
[c] Reject freq. denotes the frequency of samples with chi-squares greater than the 5 per cent critical value (expected number is 1·25).
[d] Based on a random sample of $N = 1000$.

Table 3. Chi-square for all cases[a]

Method	Case 1	Case 2	Case 3	Case 4	Case 5
GLS on y^*					
Mean	2·793	—	—	1·603	—
Variance	6·157	—	—	1·557	—
Reject freq.	2	—	—	0	—
ML					
Mean	2·525	3·123	2·540	4·644	2·300
Variance	4·684	6·363	4·978	14·296	12·533
Reject freq.	3	2	3	8	3
GLS					
Mean	2·643	3·013	2·519	4·679	2·274
Variance	5·070	5·014	4·929	15·558	11·294
Reject freq.	3	2	3	8	3
ADF					
Mean	2·492	2·706	1·840	2·633	1·894
Variance	4·363	4·829	2·576	5·430	7·359
Reject freq.	3	2	1	2	2
CVM					
Mean	—	—	—	1·527	—
Variance	—	—	—	2·357	—
Reject freq.	—	—	—	0	—

[a] Degrees of freedom = 2; Expected mean = 2; Expected variance = 4. Reject freq. denotes the frequency of samples with chi-squares greater than the 5 per cent critical value (expected number is 1·25).

regarding chi-square means, chi-square reject frequencies, parameter estimate bias, or sampling variability. Thus, we feel reasonably confident that the number of replications used in this study do not present any serious limitations.

Results will first be given pertaining to ML, GLS and ADF estimation of the y-model for case 1, case 2, case 3 and case 4, where the latter three cases exhibit an increasing skew and kurtosis. Then CVM estimation of case 4 will be presented. Finally, ML, GLS and ADF estimation of case 5 is treated.

5.1. *ML, GLS, ADF for cases 1–4 and the* $\Sigma(\mathrm{y})$ *model*

Let us first consider results for ML, GLS and ADF estimation of the y-model for cases 1, 2, 3 and 4. Table 3 gives results on chi-squares, Table 4 gives results on parameter estimates, whereas Table 5 gives results on sampling variability. Regarding chi-squares, we note that only case 4, with a univariate skew of about two and a univariate kurtosis of about three, gives substantial overestimation with ML and GLS. ADF performs rather well throughout, again keeping in mind the crudeness of the Monte Carlo study. This result on chi-square is in line with Boomsma, Tanaka and Browne.

Regarding the parameter estimates in Table 4 we find no substantial and consistent bias for any of these three estimators. For ML this is in line with Boomsma's and Browne's findings. Note, however, that this is in contrast to Tanaka's results on the non-robustness of ML and ADF and Browne's result on the non-robustness of ADF for parameter estimation with non-normal variables.

In Table 5, we find results on sampling variability for cases 1–4. A consistent bias is found for case 3 and case 4, where the ML and GLS estimated standard errors

Table 4. Parameter estimates for cases 1–5: ML, GLS and ADF[a]

Case	Parameter	True value: y Model	True value: y^* Model	Estimator: ML	Estimator: GLS	Estimator: ADF
1	λ	0·700	0·700	0·691	0·693	0·693
				(−1·3/−1·3)[b]	(−1·0/−1·0)	(−1·0/−1·0)
	θ	0·461	0·418	0·462	0·461	0·460
				(0·2/10·5)	(0·0/10·3)	(−0·2/10·0)
	ψ	0·731	0·820	0·762	0·759	0·760
				(4·2/−7·1)	(3·8/−7·4)	(4·0/−7·3)
2	λ	0·700	0·700	0·695	0·704	0·695
				(−0·7/−0·7)	(0·6/0·6)	(−0·7/−0·7)
	θ	0·767	0·694	0·772	0·769	0·768
				(0·7/11·2)	(0·3/10·8)	(0·1/10·7)
	ψ	1·210	1·360	1·225	1·210	1·227
				(1·2/−9·0)	(0·0/−11·0)	(1·4/−9·8)
3	λ	0·700	0·700	0·704	0·704	0·705
				(0·6/0·6)	(0·6/0·6)	(0·7/0·7)
	θ	0·642	0·561	0·634	0·632	0·631
				(−1·2/13·0)	(−1·6/12·7)	(−1·7/12·5)
	ψ	0·935	1·100	0·963	0·964	0·962
				(3·0/−12·5)	(3·1/−12·4)	(2·9/−12·5)
4	λ	0·700	0·700	0·704	0·706	0·701
				(0·6/0·6)	(0·9/0·9)	(0·1/0·1)
	θ	0·818	0·636	0·820	0·813	0·812
				(0·2/28·9)	(−0·6/28·6)	(−0·7/27·7)
	ψ	0·877	1·248	0·910	0·908	0·916
				(3·8/−27·1)	(3·5/−27·2)	(4·4/−26·6)
5	λ	0·700	0·700	0·696	0·696	0·696
				(−0·6/−0·6)	(−0·6/−0·6)	(−0·6/−0·6)
	θ	0·357	0·281	0·360	0·359	0·359
				(0·8/28·1)	(0·6/27·8)	(0·6/27·8)
	ψ	0·395	0·550	0·409	0·408	0·407
				(3·5/−25·6)	(3·3/−25·8)	(3·0/−26·0)

[a] Estimates are based on averaging free parameters within parameter types.
[b] In parentheses is given a/b where: a is the percentage under- or overestimation of the true value in the y model; b is the percentage under- or overestimation of the true value in the y^* model.

seem to be biased downwards as compared to the empirical standard deviations of the estimates. For case 4, this bias may be deemed serious. Regarding ML, this is in line with Boomsma, Tanaka and Browne; here the finding is generalized to GLS. However, for ADF no such bias is detected. This is in contrast with Tanaka's and Browne's studies, since they exhibited small downward biases also for ADF. We conclude that for case 4, the observations from our study are in part different from those of Tanaka and Browne. Most importantly, we find no ADF parameter estimate bias. This is an important practical point which may warrant future research. Several factors may play a role, such as: the use of ordered, categorical variables versus continuous, interval scaled ones, the degree of non-normality (skewness/kurtosis), and the type of model, including the number of degrees of freedom.

Table 5. Sampling variability for cases 1–5: ML, GLS, ADF[a]

		Estimator		
Case	Parameter	ML	GLS	ADF
1	λ	0·046	0·046	0·044
		0·044	0·044	0·044
		0·042	0·044	0·043
	θ	0·028	0·028	0·028
		0·028	0·028	0·028
		0·025	0·025	0·025
	ψ	0·073	0·073	0·076
		0·074	0·074	0·074
		0·057	0·057	0·057
2	λ	0·046	0·046	0·047
		0·046	0·046	0·046
		0·048	0·047	0·048
	θ	0·046	0·046	0·051
		0·047	0·047	0·050
		0·053	0·052	0·053
	ψ	0·121	0·121	0·117
		0·121	0·121	0·121
		0·122	0·114	0·122
3	λ	0·048	0·048	0·057
		0·047	0·047	0·053
		0·052	0·052	0·053
	θ	0·039	0·039	0·044
		0·038	0·039	0·046
		0·044	0·044	0·044
	ψ	0·098	0·098	0·108
		0·098	0·098	0·111
		0·099	0·098	0·100
4	λ	0·060	0·060	0·083
		0·061	0·060	0·083
		0·086	0·085	0·087
	θ	0·050	0·050	0·073
		0·050	0·050	0·075
		0·074	0·073	0·074
	ψ	0·110	0·110	0·155
		0·111	0·111	0·155
		0·165	0·160	0·165
5	λ	0·059	0·059	0·070
		0·059	0·059	0·071
		0·074	0·073	0·073
	θ	0·022	0·022	0·028
		0·022	0·022	0·028
		0·028	0·027	0·028
	ψ	0·048	0·048	0·062
		0·050	0·050	0·061
		0·060	0·060	0·058

[a] The three entries are:
—expected (true) standard error of estimate;
—mean of estimated standard errors calculated as the average of the standard errors for each parameter estimate;

5.2. *ML, GLS, ADF for cases 1–4 and the* $\Sigma(\mathbf{y}^*)$ *model*

Consider now parameters for the y^*-model in cases 1–4 with ML, GLS and ADF estimation. The results on chi-squares presented in Table 3 are relevant also for the y^* model, since the same structure holds. In Table 4 we find true parameter values in the column 'y^*-model'. These are the original parameter values scaled to a metric corresponding to the population y variances for each of the four cases. Here, the attenuation factor discussed in Section 2 comes into play. ML, GLS and ADF consistently estimate parameters of the y model. These parameters are related to those of the y^* model by (5), showing the expected bias of ML, GLS and ADF estimates relative to y^* model parameters. Above, we have noted no appreciable bias in ML, GLS or ADF estimation of y model parameters. Hence, the $\Sigma(\mathbf{y}^*)$ biases expressed by (5) will actually be realized in the samples as expected. Throughout, no bias is found in loadings, but is absorbed into error and factor variance bias as expected. We find that there is a consistent and increasing bias over the four cases, such that error variances are overestimated and factor variance is underestimated in line with (5). Regarding ML, this is in line with Olsson's results. We note that the bias is about the same for ML, GLS and ADF. For case 4, where the attenuation factor is 0·703, there is considerable bias of about 30 per cent. For case 1, case 2 and case 3, however, we may perhaps be willing to accept the amount of bias (generally less than 10 per cent).

Regarding sampling variability under the y^* model one may want to compare the results of Table 5 to those of Table 2 for the y^*s. Note, however, that one would then have to scale the Table 2 results to correspond to the y variances for each of the four cases. This will not be carried out here due to the parameter estimate bias already observed.

5.3. *CVM for case 4 and the* $\Sigma(\mathbf{y}^*)$ *model*

If the y^* model is of interest, the ML, GLS and ADF estimators were found to be not at all robust for case 4. Hence, for case 4, the CVM estimator was also applied, as discussed in Section 3. We then consider different association measures, the estimated latent correlations among the y^*s. Here, each y variable is dichotomized into a 25/75 per cent split. We may note in passing that the ML, GLS and ADF estimators could be applied to these 0/1 y-variables. However, this would involve an attenuation factor $a = 0{\cdot}614$ and would thus give a larger bias relative to $\Sigma(\mathbf{y}^*)$ parameters than that found in Section 5.2 ($a = 0{\cdot}703$). Results on chi-square test of fit for CVM are given at the bottom of Table 3 above. We do note a slight 'underestimation' of chi-square, but not more marked than for GLS on case 4 y^*s (see Table 2).

In Table 6 results are given for parameter estimates, and sampling variability. Here, the CVM estimator is shown to work rather well. No parameter estimate bias can be detected, and mean standard errors for parameter estimates are reasonably close to the empirical standard deviations. We note, however, a slight but consistent upward bias of mean estimated standard errors relative to the empirical standard deviations. For this situation, we conclude that the CVM estimator is a better

—empirical standard deviations of estimates calculated as the square root of the average variance for each parameter estimate.

[b] For ADF the expected standard errors of estimates are estimated from a random sample of $N = 10\,000$.

Table 6. Results for case 4 with CVM

Parameter	True value (y^* model)	Estimate[a]	Variability[b]
λ_1	0·700	0·700	—
		—	—
			—
λ_2	0·700	0·705	0·066
		(0·7)	0·065
			0·060
λ_3	0·700	0·694	0·063
		(−0·9)	0·064
			0·055
λ_4	0·700	0·711	0·066
		(1·6)	0·064
			0·055
	Factor variance		
ψ	1·000	1·003	0·123
		(0·3)	0·125
			0·110

[a] In parentheses is given the percent under- or overestimation of the true parameter value.
[b] The three entries are:
—expected (true) standard error of estimate (estimated as for ADF in Table 5);
—mean of estimated standard error;
—empirical standard deviation of estimates.

alternative than either ML, GLS or ADF. In passing, it is interesting to note a 40–50 per cent increase in sampling variability with CVM using dichotomized y variable information as compared to using y^* variable information as was the case of the GLS estimator in Table 2. Given the high precision of parameter estimation, the loss of information is however unimportant in this situation.

5.4. *ML, GLS, ADF for case 5 and* $\Sigma(\mathbf{y})$, $\Sigma(\mathbf{y}^*)$ *models*

Turning finally to case 5, we will again study the ML, GLS and ADF estimators. Here we have approximately the same univariate kurtosis as in case 4, but zero skewness. For case 4 we noted that ML and GLS estimates, but not chi-square or standard errors, were robust against the non-normality. ADF was found robust on al three counts. Will the ML/GLS problems persist when only kurtosis and not skewnes deviate from zero? Also of interest for case 5 is whether we can replicate Tanaka's findings regarding the problems of ML and ADF.

In the right-most column of Table 3 we have the case 5 chi-squares. In terms of chi-square mean and reject frequency, there seems to be no important difference between ML, GLS and ADF, and all perform close to expectation. ML and GLS robustness seems to be at hand. Parameter estimates are given in Table 4. Compared to the true parameter values of the $\Sigma(\mathbf{y})$ model, ML, GLS and ADF all perform very well with no bias or discernible differences among estimators. This result is in contrast to Tanaka's, although it should again be kept in mind that his kurtosis values were generally higher and that he did not study ordered categorical variables.

In Table 5 the sampling variability of parameter estimates is given for case 5. As with case 3 and case 4, there is a certain tendency for empirical standard deviations of ML and GLS to be larger than the other two measures of variability. Here robustness is questionable. ADF performs well with respect to these three measures of variability. In conclusion, case 5 does not replicate the findings of Tanaka in that here ADF performs well and better than both ML and GLS for $\Sigma(y)$ parameters.

It is interesting to contrast the results on ML and GLS (non-)robustness for case 4 versus case 5. Case 4 gives much more severe distortions. An interesting question that remains is the role of skewness versus kurtosis in this comparison. Case 4 does exhibit a larger skew, but it also exhibits a larger kurtosis. It should be noted that we also studied a slightly less non-normal version of case 4 with univariate skew $= -1{\cdot}921$ and univariate kurtosis $= 2{\cdot}618$, i.e. smaller kurtosis than case 5. The results were virtually identical with those of case 4 reported here. Further research is needed to evaluate the role of skewness versus kurtosis measures as for predictive value regarding distortions/non-robustness in finite samples.

Turning finally to case 5 parameter estimates as compared to true values of the $\Sigma(y^*)$ structure, we first note from Table 1 an important y correlation attenuation factor of 0·718. This results in an error variance and factor variance bias in Table 4 of 25–30 per cent. Hence, important distortions of y^* structure can occur not only in the context of skewness as was discussed in Olsson and Mooijaart, but also with zero skewness and high kurtosis.

6. Conclusions

The availability of the two new approaches ADF (for y models) and CVM (for y^* models) seems to be promising for the analysis of non-normal, ordered categorical variables. In our study, we have found that with strong skewness and/or kurtosis, these estimators may outperform the more traditional, normal theory estimators ML and GLS. Nevertheless, cost of estimation will presumably continue to be a factor of overriding concern for quite some time. From this point of view, ADF and CVM are less attractive. The approaches are roughly as costly as and considerably more costly than ML and GLS. At present, both ADF and CVM are perhaps intractable for analyses with more than, say, 25–30 variables. Hence, for still some time to come, the robustness of ML and GLS is of great interest. It is therefore reassuring to find that these normal theory estimators perform quite well even with ordered categorical and moderately skewed/kurtotic variables, at least when the sample size is not small. For our choice of $N = 1000$, no discernible differences existed between ML and GLS. From other research, however, it is conjectured that GLS may exhibit larger bias than ML in smaller samples, say $N < 400$.

If most variables have univariate skewnesses and kurtoses in the range $-1{\cdot}0$ to $+1{\cdot}0$, not much distortion is to be expected. Here ADF or CVM are not needed. From the results of Boomsma and Olsson we conjecture that this is largely independent of number of variables and number of categories. From their studies we also would not expect that our choice of equal univariate distributions would limit our conclusions. However, size of correlations seems to be important, so that with many low correlations (say 0·2 and lower), larger skews may be acceptable. Note that our study concentrates on medium-sized correlations.

When most skewnesses and/or kurtoses are larger in absolute value than 2·0, and correlations are large (say 0·5 and higher), distortions of ML and GLS chi-squares and standard errors are very likely, although estimates seem robust when relating to the y model. Unless only estimates are of interest, ML and GLS should presumably be abandoned here. ADF would be expected to still yield good chi-squares and standard errors and, if the y model is relevant, also good estimates. Note, however, that the contradictory research results of Tanaka and Browne indicate that this is specific to the type of data we have considered.

We note that ML, GLS and ADF are really intended for use with interval-scaled, unlimited variables. When, however, as in case 4, the y distribution is strongly censored, a y model is inappropriate since the assumptions of the linear factor analysis model are unrealistic. The y^* model assumed by CVM may be more appropriate.

Acknowledgements

This research was supported by Grant No. SES-8312583 from the National Science Foundation. We acknowledge the helpful assistance of Bill Cullian and Cheryl Morris. We are also grateful to M. W. Browne, G. J. Huba, L. L. Harlow, J. S. Tanaka and A. Mooijaart for providing helpful comments.

References

Bentler, P. M. (1983). Simultaneous equations as moment structure models: With an introduction to latent variable models. *Journal of Econometrics,* **22**, 13–42.

Boomsma, A. (1983). On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality. Unpublished PhD dissertation, University of Groningen, Groningen, The Netherlands.

Browne, M. W. (1974). Generalized least squares estimators in the analysis of covariance structures. *South African Statistical Journal,* **8**, 1–24.

Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (ed.), *Topics in Applied Multivariate Analysis*. London: Cambridge University Press.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology,* **37**, 62–83.

Fuller, E. L. & Hemmerle (1966). Robustness of the maximum likelihood estimation procedure in factor analysis. *Psychometrika,* **31**, 255–266.

Huba, G. J. & Bentler, P. M. (1983). Test of drug use causal model using asymptotically distribution free methods. *Journal of Drug Education,* **13**, 3–17.

Huba, G. J. & Harlow, L. L. (1983). Comparison of maximum likelihood, generalized least squares, ordinary least squares and asymptotically distribution free parameter estimates in drug abuse latent variable causal models. *Journal of Drug Education,* **13**, 387–404.

Huba, G. J. & Harlow, L. L. (1984). Robust estimation for causal models: A comparison of methods in some developmental data sets. In R. M. Lerner & D. L. Featherman (eds), *Life-span Development and Behavior,* vol. 6. New York: Academic Press.

Huba, G. J. & Tanaka, J. S. (1983). Confirmatory evidence for three daydreaming factors in the short imaginal processes inventory. *Imagination, Cognition and Personality,* **3**, 139–147.

Huba, G. J., Wingard, J. A. & Bentler, P. M. (1981). A comparison of two latent variable causal models for adolescent drug use. *Journal of Personality and Social Psychology,* **40**, 180–193.

IMSL. (1982). International Mathematical Subroutine Library.

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

Jöreskog, K. G. (1977). Structural equation models in the social sciences: Specification, estimation and testing. In P. R. Krisnaiah (ed.), *Applications of Statistics*. Amsterdam: North-Holland.

Jöreskog, K. G. & Goldberger, A. S. (1972). Factor analysis by generalized least squares. *Psychometrika,* **37**, 243–260.

Kendall, M. & Stuart, A. (1977). *The Advanced Theory of Statistics*, vol. 1, 4th ed. New York: MacMillan.

McDonald, R. P. (1974). Difficulty factors in binary data. *British Journal of Mathematical and Statistical Psychology,* **27**, 82–99.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika,* **57**, 519–530.

Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis to testing normality and to robustness studies. *Sankhya*, **36**, 115–128.
Mardia, K. V. & Zemroch, P. J. (1975). Measures of multivariate skewness and kurtosis. *Applied Statistics*, **24**, 262–265.
Mooijaart, A. (1983). Two kinds of factor analysis for ordered categorical variables. *Multivariate Behavioral Research*, **18**, 423–441.
Muthén, B. (1978). Contributions to factor analysis of dichotomous variables. *Psychometrika*, **43**, 551–560.
Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, **49**, 115–132.
Olsson, U. (1979). On the robustness of factor analysis against crude classification of the observations. *Multivariate Behavioral Research*, **14**, 485–500.
Tanaka, J. S. (1984). Some results on the estimation of covariance structure models. Unpublished PhD dissertation, University of California, Los Angeles.

Received 30 November 1984

Requests for reprints should be addressed to Bengt Muthén, Graduate School of Education, University of California, Los Angeles, CA 90024, USA.
David Kaplan is also at the above address.