PSYCHOMETRIKA—VOL. 65, NO. 4, 457–474
DECEMBER 2000

MAXIMUM LIKELIHOOD ESTIMATION OF LATENT INTERACTION EFFECTS WITH THE LMS METHOD

ANDREAS KLEIN AND HELFRIED MOOSBRUGGER

JOHANN WOLFGANG GOETHE-UNIVERSITY, FRANKFURT AM MAIN, GERMANY

In the context of structural equation modeling, a general interaction model with multiple latent interaction effects is introduced. A stochastic analysis represents the nonnormal distribution of the joint indicator vector as a finite mixture of normal distributions. The Latent Moderated Structural Equations (LMS) approach is a new method developed for the analysis of the general interaction model that utilizes the mixture distribution and provides a ML estimation of model parameters by adapting the EM algorithm. The finite sample properties and the robustness of LMS are discussed. Finally, the applicability of the new method is illustrated by an empirical example.

Key words: latent interaction effects, mixture distribution, ML estimation, structural equation modeling (SEM), EM algorithm.

1. Introduction

During the last two decades, the application of structural equation modeling (SEM) has become a common tool in the social and behavioral sciences, due to the fact that SEM integrates various statistical concepts (e.g., confirmatory factor analysis, path analysis, multiple regression, ANOVA, simultaneous equation models). The SEM approach makes it possible to analyze latent variable models, so that relationships between unobservable, latent variables can be formulated in structural equations and errors of the observed indicator variables are incorporated in measurement models. By the development of several software packages as EQS (Bentler, 1995; Bentler & Wu, 1993), LISREL (Jöreskog & Sörbom, 1993), or AMOS (Arbuckle, 1997), SEM has become available to a large community of researchers.

In a structural equation, the latent variables are usually linearly related, that is, the latent endogenous variables are linear functions of the latent exogenous variables. But in some cases theory may suggest that the effect of a latent exogenous variable on a latent endogenous variable is itself moderated by a second exogenous variable. Then, in addition to the linear effects, a latent interaction effect becomes part of the latent model structure, which means that the slope of the regression of the endogenous variable on an exogenous variable varies with the realizations of a second exogenous variable, the 'moderator variable'. The interaction effect is implemented by including a product of latent exogenous variables in the structural equation. More general, latent interaction models involve nonlinear structural relationships including one or several products of exogenous variables in the structural equation.

From the statistical standpoint, the nonlinear structural relationships have an important consequence for the distribution of variables which has been recognized (Jöreskog & Yang, 1996) but has not yet been analyzed in detail. Even if all latent exogenous variables are normally distributed, the distributions of the latent endogenous variables and their indicator variables are nonnormal. Depending on the size of the structural equation coefficients and covariances of the variables included in the nonlinear product terms, the multivariate distribution of the indicators of

This research has been supported by a grant from the Deutsche Forschungsgemeinschaft, Germany, No. Mo 474/1 and Mo 474/2. The data for the empirical example have been provided by Andreas Thiele of the University of Frankfurt, Germany. The authors are indebted to an associate editor and to three anonymous reviewers of *Psychometrika* whose comments and suggestions have been very helpful.

Requests for reprints should be sent to Andreas Klein, Department of Psychology, Mertonstrasse 17, D-60054 Frankfurt am Main, GERMANY. E-Mail: a.klein@psych.uni-frankfurt.de

0033-3123/2000-4/1998-0647-A $00.75/0

latent endogenous variables deviates substantially from normality (Moosbrugger, Schermelleh-Engel & Klein, 1997).

Ignoring the nonnormal distribution of indicator variables, the application of an estimation procedure can lead to different statistical problems: Either, if an estimation procedure is used under the assumption of multivariate normality, it must be robust against the type of nonnormality implied by latent interaction. Or, if an asymptotically distribution-free estimation method is used, it does not exploit the specific distributional characteristics of latent interaction models, which might lower the method's efficiency and power, especially when sample size is not very high.

Several researchers have called for methods to estimate latent interaction effects in structural equation models (Aiken & West, 1991; Cohen & Cohen, 1975; Jaccard, Turrisi & Wan, 1990; Schmitt, 1990). Different approaches to the estimation of latent interaction models with continuous variables have been developed. Kenny and Judd (1984) were one of the first who proposed a model for the estimation of interaction effects in latent variables. The structural equation included two latent exogenous variables and one latent product term to model the latent interaction effect. This model, henceforth called the "elementary interaction model," served as a prototype for the development of various estimation techniques.

Hayduk (1987) established the estimation of the elementary interaction model in LISREL 7-ML (Jöreskog & Sörbom, 1989), using covariance structure analysis and maximum likelihood estimation for normally distributed indicator variables. For a correct model specification, Hayduk formed a measurement model for the latent product term by multiplying indicators of exogenous variables and introduced many phantom variables and nonlinear constraints for the model implied covariance matrix. Two-step approaches for LISREL 7-ML were developed to simplify the arduous specification task of Hayduk's approach (Moosbrugger, Frank & Schermelleh-Engel, 1991; Ping, 1996a, 1996b). In the first step, Ping estimates the factor loadings and error variances of the measurement model and calculates loadings and error variances for products of indicators used as a measurement model for the latent product term. In the second step, the coefficients of the structural equation are estimated, whereas the parameters of the measurement model are fixed to their values calculated in the first step. But the application of the LISREL-ML procedures assumes multivariate normality of the indicator variables, and this assumption is violated in latent interaction models. Because products of indicators are formed to serve as indicators for the product term, the number of nonnormal variates is even more increased. In addition to violation of distributional assumptions, the two-step methods do not provide a simultaneous estimation of all model parameters.

Jöreskog and Yang (1996, 1997) and Yang Jonsson (1997) used the nonlinear capacity of LISREL 8 (Jöreskog & Sörbom, 1993) to implement the elementary interaction model. Thus, their approach avoids the multiple nonlinear constraints necessary for the phantom variables of Hayduk's method and, in contrast to the two-step approaches, establishes a simultaneous estimation of all model parameters. They propose LISREL-WLSA (weighted least squares based on the augmented moment matrix) as the asymptotically optimal method in LISREL, because it provides asymptotically correct standard errors for the estimates. As LISREL-WLSA requires large sample sizes for establishing the asymptotic properties, they suggest that LISREL-ML could be used in many cases, although the assumption of multivariate normality is violated in LISREL-ML. It could be confirmed in simulation studies (Schermelleh-Engel, Klein & Moosbrugger, 1998) that LISREL-ML estimates have no substantial bias in cases where the interaction effect is not too high and sample size is not too small. But the standard errors calculated by LISREL-WLSA or LISREL-ML often underestimate the true standard errors of the estimators, so correct testing of hypotheses cannot be expected for these methods, especially when sample size is too small.

Bollen (1995, 1996) developed a two-stage least squares (2SLS) estimation method for structural equation models. The 2SLS approach establishes a non-iterative estimation procedure which provides consistent parameter estimators and permits significance tests by calculating standard errors for the estimates. In this method the indicator variables are allowed to originate from nonnormal distributions. A simulation study with the 2SLS method confirmed relatively low bias for standard error estimates, which permits inferential statistics with acceptable Type I error. Still, the disadvantage of 2SLS lies in its low power and low efficiency, as shown by Schermelleh-Engel, Klein and Moosbrugger (1998).

General methodological problems concerning the usage and interpretation of interaction models are described by Moosbrugger, Schermelleh-Engel and Klein (1997). Moreover, they give an overview on estimation methods for latent interaction models classifying their estimation characteristics.

Klein, Moosbrugger, Schermelleh, and Frank (1997) presented an estimation procedure based on a complex analytical description of the distribution of indicator variables. The implementation of this method was computationally intensive and slow with regard to computing time. The approach was restricted to the elementary interaction model and could not be generalized to more elaborate models.

In this paper, a generalized latent interaction model and the "Latent Moderated Structural Equations" approach (LMS) is introduced[1], which implements a new ML estimation method especially developed for the distributional properties of this model. The new method is based on an analysis of the multivariate distribution of the joint indicator vector and takes the specific type of nonnormality implied by latent interaction effects explicitly into account. As a result, the joint distribution of indicator variables is represented as a finite mixture of normal distributions. Model implied mean vectors and covariance matrices of the mixture components are derived and utilized for a maximum likelihood estimation of the model parameters. The mixture density function of the joint indicator vector is explicated in LMS, and the ML estimates are computed with the EM (expectation maximation) algorithm (Dempster, Laird & Rubin, 1977; Redner & Walker, 1984), which is adapted to the mixture density. Moreover, LMS enables the estimation of standard errors by calculating the Fisher information matrix according to general ML estimation theory.

First simulation results and a comparison of LMS with LISREL and 2SLS with respect to the efficiency of the methods are given by Schermelleh-Engel, Klein and Moosbrugger (1998) for the analysis of the elementary interaction model.

The contents of this paper are as follows: In section 2, we introduce the "general interaction model" with multiple interaction effects in matrix notation form. In section 3, the nonnormal distribution of the joint indicator vector is analyzed and the model implied mean and covariance structures of the appropriate mixture density are derived. In section 4, the computation of ML parameter estimates with the EM algorithm is explained. In section 5, the finite sample properties of the LMS estimators with respect to bias, efficiency, the model difference test, and the robustness are discussed. In section 6, the analysis of an empirical data set illustrates the applicability of the LMS method for research practice.

2. The General Interaction Model

Most approaches to the analysis of latent interaction models concentrate on the type of an elementary interaction model proposed by Kenny and Judd (1984) with one latent interaction effect.

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta. \tag{1}$$

In this section, the concept of a general interaction model including multiple latent interactions is introduced, which allows not only the implementation of one interaction effect, but also the implementation of many interaction effects simultaneously[2].

[1] The LMS method is developed as part of the doctoral thesis of Andreas Klein at J.W. Goethe-University, Department of Psychology, Frankfurt, Germany.

[2] Because this paper focuses on the properties of interaction effects, we concentrate on a simplified structural equation which is restricted to a one-dimensional latent endogenous variable η.

In contrast to the elementary interaction model, the structural equation of the general interaction model is enhanced by a quadratic form to include multiple latent interaction effects. Then the structural equation of the general interaction model is

$$\eta = \alpha + \mathbf{\Gamma}\xi + \xi'\mathbf{\Omega}\xi + \zeta, \tag{2}$$

where η is a (1×1) latent endogenous variable, α is an (1×1) intercept term, ξ is a $(n \times 1)$ vector of latent exogenous variables, $\mathbf{\Gamma}$ is the $(1 \times n)$ coefficient matrix giving ξ's impact on η, $\mathbf{\Omega}$ is the $(n \times n)$ coefficient matrix giving the impact of the product terms $\xi_i\xi_j\,(i < j)$ on η, and ζ is the (1×1) disturbance variable with $E(\zeta) = 0$ and $\mathrm{Cov}(\zeta, \xi') = \mathbf{0}$. Matrix $\mathbf{\Omega}$ is assumed to be an upper triangular matrix with zeros in the diagonal:

$$\mathbf{\Omega} = \begin{pmatrix} 0 & \omega_{1,2} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \omega_{n-1,n} \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \tag{3}$$

where $\omega_{ij} = 0$ for $i \geq j$. The quadratic form $\xi'\mathbf{\Omega}\xi$ of the structural equation (2) is nonlinear in the ξ-variables and distinguishes the latent interaction model from ordinary linear SEMs. It includes product terms $\omega_{ij}\xi_i\xi_j\,(i < j)$ which model the interaction effects between pairs of ξ-variables on the dependent variable η.

It is easily seen, that in the special case of two ξ-variables and one interaction effect, the general model reduces to the following structural equation:

$$\eta = \alpha(\gamma_1\ \gamma_2)\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + (\xi_1\ \xi_2)\begin{pmatrix} 0 & \omega_{12} \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \zeta, \tag{4}$$

where ω_{12} equals the parameter γ_3 in the elementary model (see Equation (1)) described by Kenny and Judd (1984).

The latent variables of the structural equation cannot be observed directly, they are measured with error via measurement models. In the general interaction model, the measurement equations for the latent variables are

$$\mathbf{x} = \tau_{\mathbf{x}} + \mathbf{\Lambda}_{\mathbf{x}}\xi + \delta, \tag{5}$$

$$\mathbf{y} = \tau_{\mathbf{y}} + \mathbf{\Lambda}_{\mathbf{y}}\eta + \varepsilon, \tag{6}$$

where $\mathbf{x}$ is a $(q \times 1)$ vector of observed indicators of ξ, $\tau_{\mathbf{x}}$ is a $(q \times 1)$ vector of intercepts for $\mathbf{x}$, $\mathbf{\Lambda}_{\mathbf{x}}$ is a $(q \times n)$ factor loading matrix giving the impact of ξ on $\mathbf{x}$, and δ is a $(q \times 1)$ vector of measurement errors. Similarly, $\mathbf{y}$ is a $(p \times 1)$ vector of observed indicators of η, $\tau_{\mathbf{y}}$ is a $(p \times 1)$ vector of intercepts for $\mathbf{y}$, $\mathbf{\Lambda}_{\mathbf{y}}$ is a $(p \times 1)$ factor loading matrix giving the impact of η on $\mathbf{y}$, and ε is a $(p \times 1)$ vector of measurement errors. The following assumptions on the variables are made:

1. $\mathbf{x}$ is multivariate normal;
2. δ, ε are multivariate normal and mutually independent with $E(\delta) = \mathbf{0}$, $E(\varepsilon) = \mathbf{0}$,

$$\mathrm{Cov}(\delta, \xi') = \mathbf{0}, \mathrm{Cov}(\varepsilon, \xi') = \mathbf{0};$$

3. ζ is normal with $E(\zeta) = \mathbf{0}$, $\mathrm{Cov}(\zeta, \xi') = \mathbf{0}$, $\mathrm{Cov}(\zeta, \delta') = \mathbf{0}$, and $\mathrm{Cov}(\zeta, \varepsilon') = \mathbf{0}$.

The assumptions are comparable to that of Jöreskog and Yang (1996) in their analysis of interaction models with LISREL-ML. In contrast to LISREL-ML, the proposed LMS method does not assume $\mathbf{y}$ to be normally distributed, because the normality assumption of LISREL-ML is not fulfilled for interaction models. For identification of the general interaction model, it is further assumed that the latent exogenous variables $\xi_1, \ldots, \xi_n$ are scaled by n variables among the indicators $x_1, \ldots, x_q$, and that the latent endogenous variable η is scaled by one variable among the indicators $y_1, \ldots, y_p$. Each scaling variable is only influenced by a single latent variable (with factor loading set to one) and an error term. It should be noted that the model is not identified if both the latent intercept term α and the intercept vector $\tau_{\mathbf{y}}$ are free to be estimated; so at least one of them must be restricted.

3. Distribution Analysis of Indicator Variables

In this section, we investigate the nonlinear variable relationships of the general interaction model and analyze the type of nonnormal distribution induced by latent interaction effects. As a result, it is shown that the distribution of the joint indicator vector $(\mathbf{x}, \mathbf{y})$ can be represented as a finite mixture of multivariate normal distributions. To achieve this representation, the ξ-variables are sorted for a separation of linear and nonlinear relationships and decomposed into mutually independent random variables $z_1, \ldots, z_n$. The vector $\mathbf{z} = (z_1, \ldots, z_n)$ is partitioned into vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ in order to separate linear and nonlinear parts of the structural- and measurement equations. The subvector $\mathbf{z}_1$ is used to form an augmented random vector $(\mathbf{z}_1, \mathbf{x}, \mathbf{y})$, and the density of the augmented vector is derived. Then, the density of $(\mathbf{x}, \mathbf{y})$, which becomes the marginal density of the augmented vector, can be expressed as a continuous mixture of normal densities. The model implied mean vectors and covariance matrices of the mixture components are derived in matrix notation form. Finally, the continuous mixture density developed for $(\mathbf{x}, \mathbf{y})$ is approximated by a finite mixture of normal densities. Hermite-Gaussian quadrature formulas of numerical integration are applied to calculate the appropriate mixture probabilities and mixture components for the finite mixture distribution.

3.1. Nonnormality of η and Joint Indicator Vector $(\mathbf{x}, \mathbf{y})$

In the measurement equation of the ξ-variables (5), the indicator vector $\mathbf{x}$ and the vector δ of measurement errors are assumed to be multivariate normally distributed. Then it follows from the equations for the scaling variables among the $x_1, \ldots, x_q$ that ξ is multivariate normal. But if, under this assumption, interaction effects exist, then η is *nonnormally* distributed, because the structural equation of η (Equation (2)) contains the quadratic form $\xi'\mathbf{\Omega}\xi$. Since η is measured by the indicator vector $\mathbf{y}$, the indicator vector $\mathbf{y}$ is nonnormally distributed either. Therefore, the joint indicator vector $(\mathbf{x}, \mathbf{y})$ is multivariate *nonnormally* distributed.

All model parameters are summarized in a parameter vector θ. In order to develop an ML estimation method for parameter vector θ, the distribution of the $(q+p) \times 1$ joint indicator vector $(\mathbf{x}, \mathbf{y}) = (x_1, \ldots, x_q, y_1, \ldots, y_p)$ is analyzed in order to derive its multivariate density function. The characteristics of this nonnormal distribution have not yet been explored.

3.2. Sorting of ξ-Variables for Separation of Linear and Nonlinear Relationships

The distribution analysis of the joint indicator vector $(\mathbf{x}, \mathbf{y})$ starts from the equations of the general interaction model (i.e., (2), (5), (6)). The researcher is usually not interested in estimating all possible interaction effects between pairs of latent exogenous variables. Therefore, some elements of the upper triangle of matrix $\mathbf{\Omega}$ (see (3)) can be set to zero. Taking this into account, vector ξ is sorted so that the nonzero elements of $\mathbf{\Omega}$ appear in its first k rows only and each of the first k rows contains at least one nonzero element:

$$\boldsymbol{\Omega} = \begin{pmatrix} 0 & \omega_{1,2} & \cdots & \cdots & \cdots & \omega_{1,n} \\ \vdots & \ddots & \ddots & & & \vdots \\ 0 & \cdots & 0 & \omega_{k,k+1} & \cdots & \omega_{k,n} \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{7}$$

Without loss of generality, vector ξ can always be sorted so that $\boldsymbol{\Omega}$ has the appropriate form. Because after the sorting of vector ξ matrix $\boldsymbol{\Omega}$ has nonzero elements in its first k rows only, we have $i \leq k$ for the product terms $\omega_{ij}\xi_i\xi_j$ occurring in (2). Hence, the latent endogenous variable η is *linearly* related to $\xi_{k+1}, \ldots, \xi_n$, but can be *nonlinearly* related to $\xi_1, \ldots, \xi_k$.

3.3. Decomposition of ξ-Variables into Independent z-Variables

Let $\boldsymbol{\Phi} = \mathbf{AA}'$ be the Cholesky decomposition of the positive definite $(n \times n)$ covariance matrix $\boldsymbol{\Phi}$ of the normally distributed vector ξ, where $\mathbf{A}$ is a $(n \times n)$ lower triangular matrix. With $\mathbf{A}$ assumed to be a lower triangular matrix, there is a one-to-one correspondence between positive definite covariance matrices $\boldsymbol{\Phi}$ and matrices $\mathbf{A}$. Thus, for ML parameter estimation it is equivalent to estimate $\mathbf{A}$ instead of $\boldsymbol{\Phi}$. $\mathbf{A}$ is used for a decomposition of the ξ-variables into n mutually independent variables $z_1, \ldots, z_n$

$$\xi = \mathbf{Az}, \tag{8}$$

where $\mathbf{z} = (z_1, \ldots, z_n)'$ is a $(n \times 1)$ standardized normally distributed random vector. The vectors ξ and $\mathbf{Az}$ are identically distributed.

3.4. Partitioning of z-Variables for Separation of Linear and Nonlinear Relationships

Vector $\mathbf{z}$ can be partitioned so that, with regard to the variables $z_1, \ldots, z_n$ of $\mathbf{z}$, the linear and nonlinear parts of the structural- and measurement equations can be separated. For this purpose, a partitioned vector $\mathbf{z}$ is created, where $\mathbf{z_1} = (z_1, \ldots, z_k)'$ and $\mathbf{z}_2 = (z_{k+1}, \ldots, z_n)'$:

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{z} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{z}_2 \end{bmatrix}. \tag{9}$$

The exogenous vector η in the structural equation (2) is substituted by $\mathbf{Az}$ (i.e. (8)), the resulting expression is expanded for $\mathbf{z_1}$, $\mathbf{z_2}$ (with (9)), and the terms of $\mathbf{z_2}$ are collected:

$$\begin{aligned} \eta &= \alpha + \boldsymbol{\Gamma}\mathbf{Az} + \mathbf{z}'\mathbf{A}'\boldsymbol{\Omega}\mathbf{Az} + \zeta \\ &= \alpha + \boldsymbol{\Gamma}\mathbf{A}\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix} + \boldsymbol{\Gamma}\mathbf{A}\begin{bmatrix} \mathbf{0} \\ \mathbf{z}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}' \mathbf{A}'\boldsymbol{\Omega}\mathbf{A}\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}' \mathbf{A}'\boldsymbol{\Omega}\mathbf{A}\begin{bmatrix} \mathbf{0} \\ \mathbf{z}_2 \end{bmatrix} \\ &\quad + \begin{bmatrix} \mathbf{0} \\ \mathbf{z}_2 \end{bmatrix}' \mathbf{A}'\boldsymbol{\Omega}\mathbf{A}\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{z}_2 \end{bmatrix}' \mathbf{A}'\boldsymbol{\Omega}\mathbf{A}\begin{bmatrix} \mathbf{0} \\ \mathbf{z}_2 \end{bmatrix} + \zeta \\ &= \left(\alpha + \boldsymbol{\Gamma}\mathbf{A}\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}' \mathbf{A}'\boldsymbol{\Omega}\mathbf{A}\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}\right) + \left(\boldsymbol{\Gamma}\mathbf{A} + \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}' \mathbf{A}'\boldsymbol{\Omega}\mathbf{A}\right)\begin{bmatrix} \mathbf{0} \\ \mathbf{z}_2 \end{bmatrix} + \zeta. \end{aligned} \tag{10}$$

The last identity holds because $\begin{bmatrix} \mathbf{0} \\ \mathbf{z}_2 \end{bmatrix}' \mathbf{A}'\boldsymbol{\Omega}$ equals the zero matrix. But $\boldsymbol{\Omega}$ is not symmetric and $\boldsymbol{\Omega}\mathbf{A}\begin{bmatrix} \mathbf{0} \\ \mathbf{z}_2 \end{bmatrix}$ is nonzero. The last line of (10) shows that η is linear in $\mathbf{z_2}$, but nonlinear in $\mathbf{z_1}$, which implies that η is nonnormally distributed. For the measurement equations, the substitutions given by (8), (9), and (10) yield

$$\begin{aligned}
\mathbf{x} &= \tau_{\mathbf{x}} + \mathbf{\Lambda}_{\mathbf{x}}\xi + \delta \\
&= \tau_{\mathbf{x}} + \mathbf{\Lambda}_{\mathbf{x}}\mathbf{A}\mathbf{z} + \delta \qquad (11)\\
&= \tau_{\mathbf{x}} + \mathbf{\Lambda}_{\mathbf{x}}\mathbf{A}\begin{bmatrix}\mathbf{z}_1\\ \mathbf{0}\end{bmatrix} + \mathbf{\Lambda}_{\mathbf{x}}\mathbf{A}\begin{bmatrix}\mathbf{0}\\ \mathbf{z}_2\end{bmatrix} + \delta,\\
\mathbf{y} &= \tau_{\mathbf{y}} + \mathbf{\Lambda}_{\mathbf{y}}\eta + \varepsilon \qquad (12)\\
&= \tau_{\mathbf{y}} + \mathbf{\Lambda}_{\mathbf{y}}\left(\alpha + \mathbf{\Gamma}\mathbf{A}\begin{bmatrix}\mathbf{z}_1\\ \mathbf{0}\end{bmatrix} + \begin{bmatrix}\mathbf{z}_1\\ \mathbf{0}\end{bmatrix}'\mathbf{A}'\mathbf{\Omega}\mathbf{A}\begin{bmatrix}\mathbf{z}_1\\ \mathbf{0}\end{bmatrix}\right) + \mathbf{\Lambda}_{\mathbf{y}}\left(\mathbf{\Gamma}\mathbf{A} + \begin{bmatrix}\mathbf{z}_1\\ \mathbf{0}\end{bmatrix}'\mathbf{A}'\mathbf{\Omega}\mathbf{A}\right)\begin{bmatrix}\mathbf{0}\\ \mathbf{z}_2\end{bmatrix} + \mathbf{\Lambda}_{\mathbf{y}}\zeta + \varepsilon.
\end{aligned}$$

Equation (11) shows that $\mathbf{x}$ is linear in $\mathbf{z}_1$ and $\mathbf{z}_2$, whereas (12) shows that $\mathbf{y}$ is *linear* in $\mathbf{z}_2$, but *nonlinear* in $\mathbf{z}_1$.

3.5. Derivation of Continuous Mixture Density of Joint Indicator Vector

Vector $\mathbf{z}_1$ can be used to analyze the distribution of the joint indicator vector $(\mathbf{x}, \mathbf{y})$. For the multivariate normal distribution of $\mathbf{z}_1$, we have that

$$\mathbf{z}_1 \sim N(\mathbf{0}, \mathbf{I}_k), \qquad (13)$$

where $\mathbf{I}_k$ is the $(k \times k)$ identity matrix. Because $\mathbf{y}$ is *nonlinear* in $\mathbf{z}_1$, the distribution of $\mathbf{y}$ is nonnormal. The joint indicator vector $(\mathbf{x}, \mathbf{y})$ is also *nonlinear* in $\mathbf{z}_1$ and therefore nonnormally distributed, but it is linear in $\mathbf{z}_2$. Therefore, if values for $\mathbf{z}_1$ are given, the joint indicator vector $(\mathbf{x}, \mathbf{y})$ is *linear* in $\mathbf{z}_2$, that is, a linear combination of the normally distributed variables $z_{k+1}, \ldots, z_n$. Thus, the conditional distribution of $(\mathbf{x}, \mathbf{y}) \mid \mathbf{z}_1$ is multivariate normal:

$$(\mathbf{x}, \mathbf{y}) \mid \mathbf{z}_1 \sim N(\mu(\mathbf{z}_1), \Sigma(\mathbf{z}_1)), \qquad (14)$$

where $\mu(\mathbf{z}_1)$ denotes the $(q+p)\times 1$ model implied mean vector and $\Sigma(\mathbf{z}_1)$ denotes the $(q+p)\times(q+p)$ model implied covariance matrix of the conditioned joint indicator vector $(\mathbf{x}, \mathbf{y}) \mid \mathbf{z}_1$. They are functions of $(k \times 1)$ vector $\mathbf{z}_1$ and the model parameters θ.

If the subvector $\mathbf{z}_1$ is used to form the augmented vector $(\mathbf{z}_1, \mathbf{x}, \mathbf{y})$, the density function of $(\mathbf{x}, \mathbf{y})$ can be derived. Following directly from the definition of conditional distributions, the distribution of $(\mathbf{z}_1, \mathbf{x}, \mathbf{y})$ is the product of the distribution of $\mathbf{z}_1$ (i.e., (13)) and the distribution of $(\mathbf{x}, \mathbf{y}) \mid \mathbf{z}_1$ (i.e., (14)). The distribution of the joint indicator vector $(\mathbf{x}, \mathbf{y})$ equals the marginal distribution over $\mathbf{z}_1 = (z_1, \ldots, z_k)'$ and the resulting marginal density function f for $(\mathbf{x}, \mathbf{y})$ can be expressed as an integral over the k-dimensional real space $\mathbf{R}^k$. The density value for the realization $(\mathbf{x} = x, \mathbf{y} = y)$ is

$$\mathrm{f}(\mathbf{x} = x, \mathbf{y} = y) = \int_{\mathbf{R}^k} \varphi_{0,\mathbf{I}_k}(\mathbf{z}_1)\varphi_{\mu(\mathbf{z}_1),\Sigma(\mathbf{z}_1)}(x, y)\, d\mathbf{z}_1. \qquad (15)$$

The integral describes a continuous mixture of $(q + p)$-dimensional normal densities $\varphi_{\mu(\mathbf{z}_1)}, \Sigma_{(\mathbf{z}_1)}$ with $\mathbf{z}_1$ as $(k \times 1)$ mixing vector.

3.6. Model Implied Mean Vectors and Covariance Matrices of the Mixture Components

In order to exploit (15) for a distribution analysis of the indicator variables, it is necessary to explicate the model implied mean vector $\mu(\mathbf{z}_1)$ and covariance matrix $\Sigma(\mathbf{z}_1)$ of the conditioned joint indicator vector $(\mathbf{x}, \mathbf{y}) \mid \mathbf{z}_1$ as functions of vector $\mathbf{z}_1$. For notational purposes, mean vector and covariance matrix are partitioned:

$$\mu(\mathbf{z}_1) = \begin{bmatrix} \mu_{\mathbf{x}}(\mathbf{z}_1) \\ \mu_{\mathbf{y}}(\mathbf{z}_1) \end{bmatrix} \tag{16}$$

$$\Sigma(\mathbf{z}_1) = \begin{bmatrix} \Sigma_{\mathbf{xx}}(\mathbf{z}_1) & \Sigma_{\mathbf{xy}}(\mathbf{z}_1) \\ \Sigma'_{\mathbf{xy}}(\mathbf{z}_1) & \Sigma_{\mathbf{yy}}(\mathbf{z}_1) \end{bmatrix}, \tag{17}$$

where $\mu_{\mathbf{x}}(\mathbf{z}_1)(q \times 1)$ and $\mu_{\mathbf{y}}(\mathbf{z}_1)(p \times 1)$ denote the model implied conditioned mean vectors of $\mathbf{x} \mid \mathbf{z}_1$ and $\mathbf{y} \mid \mathbf{z}_1$, respectively. The matrices $\Sigma_{\mathbf{xx}}(\mathbf{z}_1)(q \times q)$, $\Sigma_{\mathbf{xy}}(\mathbf{z}_1)(q \times p)$, and $\Sigma_{\mathbf{yy}}(\mathbf{z}_1)(p \times p)$ denote the model implied conditioned covariance matrices of $\mathbf{x}$ given $\mathbf{z}_1$ and $\mathbf{y}$ given $\mathbf{z}_1$ in the appropriate order.

Because $(n \times 1)$ vector $\mathbf{z}$ is assumed to be standardized normally distributed, the subvectors and submatrices of the model implied mean vector and covariance matrix can now be explicated by using (11) and (12)

$$\mu_{\mathbf{x}}(\mathbf{z}_1) = \tau_{\mathbf{x}} + \mathbf{\Lambda}_{\mathbf{x}}\mathbf{A}\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}, \tag{18}$$

$$\mu_{\mathbf{y}}(\mathbf{z}_1) = \tau_{\mathbf{y}} + \mathbf{\Lambda}_{\mathbf{y}}\left(\alpha + \mathbf{\Gamma}\mathbf{A}\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}' \mathbf{A}'\mathbf{\Omega}\mathbf{A}\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}\right), \tag{19}$$

$$\Sigma_{\mathbf{xx}}(\mathbf{z}_1) = \mathbf{\Lambda}_{\mathbf{x}}\mathbf{A}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-k} \end{pmatrix}\mathbf{A}'\mathbf{\Lambda}'_{\mathbf{x}} + \mathbf{\Theta}_{\delta}, \tag{20}$$

$$\Sigma_{\mathbf{xy}}(\mathbf{z}_1) = \mathbf{\Lambda}_{\mathbf{x}}\mathbf{A}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-k} \end{pmatrix}\left(\mathbf{\Gamma}\mathbf{A} + \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}' \mathbf{A}'\mathbf{\Omega}\mathbf{A}\right)' \mathbf{\Lambda}'_{\mathbf{y}} = \Sigma'_{\mathbf{yx}}(\mathbf{z}_1), \tag{21}$$

$$\begin{aligned}\Sigma_{\mathbf{yy}}(\mathbf{z}_1) = {} & \mathbf{\Lambda}_{\mathbf{y}}\left(\mathbf{\Gamma}\mathbf{A} + \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}' \mathbf{A}'\mathbf{\Omega}\mathbf{A}\right)\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-k} \end{pmatrix}\left(\mathbf{\Gamma}\mathbf{A} + \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{0} \end{bmatrix}' \mathbf{A}'\mathbf{\Omega}\mathbf{A}\right)' \mathbf{\Lambda}'_{\mathbf{y}} \\ & + \mathbf{\Lambda}_{\mathbf{y}}\Psi_{11}\mathbf{\Lambda}'_{\mathbf{y}} + \mathbf{\Theta}_{\varepsilon},\end{aligned} \tag{22}$$

where the matrices $\mathbf{\Theta}_{\delta}$ and $\mathbf{\Theta}_{\varepsilon}$ denote the covariance matrices of the error vectors δ and ε, respectively. Ψ_{11} denotes the variance of disturbance term ζ.

Summarizing the dependencies of the subvectors and submatrices with regard to $\mathbf{z}_1$, it can be stated that submatrix $\Sigma_{\mathbf{xx}}(\mathbf{z}_1)$ is independent of $\mathbf{z}_1$, whereas $\mu_{\mathbf{x}}(\mathbf{z}_1)$ and $\Sigma_{\mathbf{xy}}(\mathbf{z}_1)$ depend on the variables $z_1, \ldots, z_k$ of $\mathbf{z}_1$ *linearly*. In case of interaction, matrix $\mathbf{\Omega}$ is different from zero matrix, and subvector $\mu_{\mathbf{y}}(\mathbf{z}_1)$ and submatrix $\Sigma_{\mathbf{yy}}(\mathbf{z}_1)$ depend on the variables $z_1, \ldots, z_k$ of $\mathbf{z}_1$ *nonlinearly*.

For the case of the elementary interaction model, the model implied mean vectors and covariance matrices are explicitly given by Schermelleh-Engel, Klein and Moosbrugger (1998).

3.7. Approximation of Continuous Mixture Density by Hermite-Gaussian Quadrature Formulas

Because of the nonlinear relationship mentioned above, the integral of the mixture density (15) *cannot* be solved analytically. Instead, the k-dimensional integral of (15) can be approximated by numerical methods, for example, by application of Hermite-Gaussian quadrature formulas for numerical integration (Isaacson & Keller, 1966). To apply Hermite-Gaussian quadrature formulas, the integration variable $\mathbf{z}_1$ is substituted by

$$\mathbf{u} = 2^{-1/2}\mathbf{z}_1, \tag{23}$$

where $\mathbf{u} = (\mathrm{u}_1, \ldots, \mathrm{u}_k)'$. Then, after inserting the k-dimensional standard normal density for $\varphi_{0,\mathbf{I}_k}(\mathbf{z}_1)$ and substituting $\mathbf{z}_1$, (15) changes to the following integral

$$\mathrm{f}(\mathbf{x} = x, \mathbf{y} = y) = \int_{\mathbf{R}^k} \pi^{-k/2} \exp(-\mathbf{u}'\mathbf{u}) \varphi_{\mu(2^{1/2}\mathbf{u}),\Sigma(2^{1/2}\mathbf{u})}(x, y)\, d\mathbf{u} \tag{24}$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp(-\mathrm{u}_1^2) \cdots \exp(-\mathrm{u}_k^2) \pi^{-k/2} \varphi_{\mu(2^{1/2}\mathbf{u}),\Sigma(2^{1/2}\mathbf{u})}(x, y)\, d\mathrm{u}_1 \cdots d\mathrm{u}_k. \tag{25}$$

The k-dimensional integral of (25) is approximated by iterated application (k times) of the Hermite-Gaussian quadrature formula, which approximates one-dimensional integrals with weight function $\exp(-\mathrm{u}^2)$ in the integrand by finite sums (Isaacson & Keller, 1966). The quadrature formula provides M weights $\mathrm{w}_j (j = 1, \ldots, M)$ and $M (k \times 1)$ node vectors $\nu_j (j = 1, \ldots, M)$ so that the density function f is approximated by a finite sum

$$\mathrm{f}(\mathbf{x} = x, \mathbf{y} = y) \approx \sum_{j=1}^{M} \mathrm{w}_j \pi^{-k/2} \varphi_{\mu(2^{1/2}\nu_j),\Sigma(2^{1/2}\nu_j)}(x, y), \tag{26}$$

where the choice of M determines the degree of exactness selected for the approximation. The weights and node vectors are derived analytically by quadrature theory for a choice of M and k. They are calculated independently of the model parameters and need not to be estimated (for lists of weights and nodes see Abramowitz & Stegun, 1971). The higher M is chosen for the quadrature, the more exact the approximation is. The accuracy of the approximation is not influenced by the model parameters.

As a result, the finite sum (26) approximates the density function f of the joint indicator vector $(\mathbf{x}, \mathbf{y})$ as a finite mixture of M normal densities. Since the weights fulfil the condition

$$\sum_{j=1}^{M} \mathrm{w}_j \pi^{-k/2} = 1, \tag{27}$$

the terms $\mathrm{w}_j \pi^{-k/2} (j = 1, \ldots, M)$ can be defined as mixture probabilities $\rho_j (j = 1, \ldots, M)$:

$$\rho_j = \mathrm{w}_j \pi^{-k/2} (j = 1, \ldots, M). \tag{28}$$

Using this definition, the expression for the density function (26) changes to the standard notation form for a finite mixture density:

$$\mathrm{f}(\mathbf{x} = x, \mathbf{y} = y) \approx \sum_{j=1}^{M} \rho_j \varphi_{\mu(2^{1/2}\nu_j),\Sigma(2^{1/2}\nu_j)}(x, y). \tag{29}$$

For the case of the elementary interaction model, the mixture density is derived by Schermelleh-Engel, Klein and Moosbrugger (1998) as an example. In this case we have $k = 1$, and the choice $M = 16$ provides a sufficiently precise approximation. For $k = 2$, a choice of $M = 24$ node vectors provided a close approximation in simulation experiments. The higher k is, the more node vectors must be used for the approximation. Although interaction models with values of k greater than 2 have not yet been tested, k should be limited in size to about 4 or 5 for the implementation of LMS on present personal computers (e.g., Pentium 200 MHz). A more complete evaluation of this awaits further analytical research.

4. ML Estimation

A common approach to the ML estimation of parametric models with mixture densities is the application of the EM (expectation maximation) algorithm (Redner & Walker, 1984). Under fairly general conditions, the EM algorithm provides an iterative estimation procedure that converges to a maximum likelihood estimation of the model parameters (Dempster, Laird & Rubin, 1977).

In the following, the adaptation of the EM algorithm for LMS is described. All model parameters are collected in the parameter vector θ, and $\theta^{(0)}$ denotes the vector of starting values for the iterative algorithm. At the beginning of step r, the current value for parameter vector θ is $\theta^{(r-1)}$, and the new value $\theta^{(r)}$ is to be calculated.

In the density function f (Equation 29), the parameter relations of the mean vectors and covariance matrices in the M mixture components are *not independent* of each other, which involves a numerically more complicated situation for the EM algorithm than finite mixtures with *independent* components.

M denotes the number of mixture components selected for the approximation of the continuous mixture density function f, and N denotes the sample size. Every iteration step consists of two subordinate steps, the expectation step and the maximation step. The *expectation step* is the calculation of the elements $\mathrm{p}^{(r)}(\mathrm{j} = j \mid x_i, y_i)$ $(i = 1, \ldots, N;\ j = 1, \ldots, M)$ of a $(N \times M)$ matrix $\mathbf{P}^{(\mathrm{r})}$ containing the posterior probabilities for the mixture components j given the i-th row of the data matrix of joint indicator vector $(\mathbf{x}, \mathbf{y})$, where $(x_i, y_i) = (x_{i1}, \ldots, x_{iq}, y_{i1}, \ldots, y_{ip})$:

$$\mathrm{p}^{(r)}(\mathrm{j} = j \mid x_i, y_i) = \frac{\rho_j \varphi_{\mu(2^{1/2}\nu_j), \Sigma(2^{1/2}\nu_j)}(x_i, y_i)}{\mathrm{f}(\mathbf{x} = x_i, \mathbf{y} = y_i)}. \tag{30}$$

A general derivation of (30) is given by Grim (1982). The numerator of Equation 30 is the density value of the j-th component of the finite mixture (29), and the denominator is the density value of (x_i, y_i). For calculation of these values, the parameter vector $\theta^{(r-1)}$ is used for computing the mean vectors $\mu(2^{1/2}\nu_j)$ and covariance matrices $\Sigma(2^{1/2}\nu_j)$ with the argument values of the node vectors $2^{1/2}\nu_j$ inserted for $\mathbf{z_1}$ (see (18) to (22)). In contrast to covariance structure analysis, here LMS uses the full information of the raw data. The weights ρ_j and node ν_j vectors are derived analytically and need not to be estimated.[3]

The *maximation step* is the calculation of the new parameter vector $\theta^{(r)}$ as the argument value of θ which maximizes a sum weighted by the posterior probabilities computed in the expectation step:

$$\theta^{(r)} = \arg\max_{\theta} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{M} \mathrm{p}^{(r)}(\mathrm{j} = j \mid x_i, y_i) \ln \varphi_{\mu(2^{1/2}\nu_j), \Sigma(2^{1/2}\nu_j)}(x_i, y_i) \right\}, \tag{31}$$

where $\mu(2^{1/2}\nu_j)$ and $\Sigma(2^{1/2}\nu_j)$ $(j = 1, \ldots, M)$ on the right-hand side of (31) are functions of free parameter vector θ. A general derivation of Equation 31 is given by Grim (1982). For the calculation of $\theta^{(r)}$, numerical computation of partial derivatives and a straightforward application of the single step iteration method (Isaacson & Keller, 1966; Schwarz, 1993) for solving systems of nonlinear equations are used in LMS. The sequence $[\theta^{(r)}]_{r=0,1,2,\ldots}$ provided by iteration of the EM steps converges to a ML estimation $\hat{\theta}$ of θ.

For a flow chart which illustrates the different steps of the EM estimation procedure, see Moosbrugger, Schermelleh-Engel and Klein (1997). The EM algorithm is a computationally intensive procedure, and the parameter estimation with LMS (algorithm programmed in Delphi program code) on a personal computer (Pentium 200 MHz) typically needs from 10 to 60 EM iteration steps and from 5 to 60 seconds, depending on the model, sample size and starting values. The calculation of the standard errors is more complex and can require up to several minutes of computing time.

[3]In a standard mixture analysis with unknown mixing weights, the average posterior probabilities equal the estimated mixing weights after the convergence of the EM algorithm, that is, the equality $\sum_i^N \mathrm{p}(\mathrm{j} = j \mid x_i, y_i)/N = \hat{\rho}_j$ holds. The equality is asymptotically true for the mixture analysis described here, where $\hat{\rho}_j$ is replaced by the analytically derived mixing weight ρ_j.

5. Properties of LMS Estimators

In LMS, the density function of joint indicator vector $(\mathbf{x}, \mathbf{y})$ is represented as a finite mixture of normal densities, and LMS utilizes the model implied mean vectors and covariance matrices of the mixture components for an iterative estimation of the model parameters with the EM algorithm. Based on the distribution analysis of the joint indicator vector, LMS takes the non-normality of the distribution explicitly into account. It provides ML estimators, and their large sample properties are given by general ML estimation theory: they are consistent, asymptotically unbiased, asymptotically efficient, and asymptotically normally distributed. The standard errors for the LMS estimates can be computed from the Fisher information matrix, as follows from general ML estimation theory. For the elements of the Fisher information matrix, partial derivatives of the logarithm of the density function are used, and expectation values of this function are calculated. The expectation values are computed by using a simulated large sample which follows the distribution of the indicator vector given by the LMS parameter estimates.

Moreover, again based on the density analysis of the general interaction model, one can carry out a model difference test for interaction hypotheses by calculating the likelihood ratio test statistic. In a model difference test, the loglikelihood for the interaction model (with free parameters in $\mathbf{\Omega}$) is compared to the loglikelihood of a more restricted model (e.g., with all parameters in $\mathbf{\Omega}$ set to zero).

For the analysis of empirical data, LMS assumes that the specified interaction model holds and that the indicator vector $\mathbf{x}$ of the latent exogenous vector ξ is multivariate normally distributed, which should be verified before using LMS. Unlike covariance structure analysis, LMS uses the raw data of indicator variables directly for estimation, *and does not require the forming of any products of indicator variables.*

5.1. Bias and Efficiency

The finite sample properties of LMS estimators have been examined in a Monte-Carlo study by Schermelleh-Engel, Klein and Moosbrugger (1998), where LMS was compared to three other estimation methods: LISREL-WLSA, LISREL-ML (Jöreskog & Yang, 1996, 1997; Yang Jonsson, 1997), and 2SLS (Bollen, 1995, 1996). In the study, sample size and interaction effect size were varied at different levels in order to test the performance of the methods under different conditions. In this section, simulation results for the analysis of an elementary interaction model (4) with the following measurement equations are reported:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda_{x_{21}} & 0 \\ 0 & 1 \\ 0 & \lambda_{x_{42}} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{pmatrix}, \tag{32}$$

$$y = \eta. \tag{33}$$

The elementary interaction model has 14 model parameters, and for data generation their values (see Table 1, column True Value) were taken from an example of Jöreskog and Yang (1996). Using the PRELIS program (Jöreskog & Sörbom, 1996), 500 data sets of sample size $N = 400$ for the five indicator variables x_1, x_2, x_3, x_4, y were generated. The data were generated with intercept vectors τ_x and τ_y set to zero. LMS, LISREL-WLSA, LISREL-ML, and 2SLS analyzed each data set separately, computing 500 estimates for every model parameter (with the exception of 2SLS, where only the parameters of the structural equation $(\alpha, \gamma_1, \gamma_2, \omega_{12})$ were estimated). For the analysis of the model, intercept vectors τ_x and τ_y did not need to be estimated. In the study, all methods provided unbiased parameter estimates, as the means of the estimates over all 500 data sets showed no substantial deviation from the true parameter values. Therefore, the means of the estimates are not reported here. The efficiency of the parameter estimators was examined by calculation of standard deviations (MC-SDs) of the distributions of estimates

TABLE 1.
Estimation results of a Monte-Carlo study for the elementary interaction model with one latent interaction effect (4). 500 data sets of sample size $N = 400$ were analyzed with LMS, LISREL-WLSA, LISREL-ML, and 2SLS. The columns give for every model parameter: the true value, the standard deviation of Monte Carlo parameter estimates (MC-SD), and the mean of estimated standard errors (Est-SE) over all 500 data sets. For the Hermite-Gaussian quadrature formula (26) $M = 16$ was chosen.

	True	LMS		LISREL-WLSA		LISREL-ML		2SLS	
Parameter	Value	MC-SD	Est-SE	MC-SD	Est-SE	MC-SD	Est-SE	MC-SD	Est-SE
α	1.000	0.032	0.033	0.044	0.037	0.042	0.038	0.067	0.066
γ_1	0.200	0.064	0.065	0.091	0.075	0.089	0.076	0.156	0.165
γ_2	0.400	0.061	0.061	0.079	0.065	0.076	0.062	0.114	0.115
ω_{12}	0.700	0.094	0.102	0.155	0.107	0.161	0.112	0.255	0.221
Ψ_{11}	0.200	0.025	0.024	0.094	0.101	0.050	0.039	—	—
$\lambda_{x_{21}}$	0.600	0.092	0.099	0.136	0.085	0.121	0.079	—	—
$\lambda_{x_{42}}$	0.700	0.077	0.078	0.100	0.067	0.092	0.060	—	—
ϕ_{11}	0.490	0.077	0.081	0.140	0.086	0.129	0.078	—	—
ϕ_{21}	0.235	0.044	0.040	0.058	0.032	0.049	0.031	—	—
ϕ_{22}	0.640	0.087	0.086	0.115	0.074	0.110	0.068	—	—
$\theta_{\delta_{11}}$	0.510	0.059	0.061	0.126	0.099	0.109	0.075	—	—
$\theta_{\delta_{22}}$	0.640	0.050	0.054	0.091	0.068	0.060	0.039	—	—
$\theta_{\delta_{33}}$	0.360	0.060	0.057	0.093	0.081	0.077	0.058	—	—
$\theta_{\delta_{44}}$	0.510	0.046	0.047	0.078	0.067	0.052	0.036	—	—

(see Table 1). In the study, the standard deviations of the estimates (MC-SD) indicated a higher efficiency for the LMS parameter estimators when compared to the alternative methods.

For every data set, each method estimates also the standard errors for the parameter estimates. The means of these estimated standard errors (Est-SE) were computed and compared to the standard deviations of the estimates (MC-SD). In LMS, the means of these estimated standard errors (Est-SE), calculated over all 500 data sets, are very close to the standard deviations calculated from the estimates (MC-SD). Thus, the Monte-Carlo study did not reveal any substantial bias for the LMS estimation of standard errors. Therefore, when LMS analyzed a single data set, the estimated standard errors for the parameter estimates were reliable measures which could be used for calculating confidence intervals and for testing of hypotheses about the parameters.

The LISREL-WLSA method tested in the study is asymptotically distribution free, so there are no distributional assumptions violated when interaction models are analyzed. LISREL-WLSA provides consistent parameter estimates, but despite of these theoretical properties, the application of this method is limited in practice: The results of the study showed that the efficiency of the LISREL-WLSA estimates is low compared to LMS (see Table 1). Further, when Est-SE and MC-SD are compared for LISREL-WLSA, the standard errors are seriously underestimated for some parameters.

The LISREL-ML approach provided consistent and unbiased parameter estimates. Although LISREL-ML requires normally distributed indicator variables, an assumption which is generally violated in interaction models, it could be used in cases where the interaction effect is not too high and sample size is not too small. The means of the standard errors given in LISREL-ML outputs (Est-SE) often underestimate the Monte Carlo standard deviation (MC-SD) in Table 1. The relative bias of the standard errors rises with increasing interaction effect because the normality assumption for the indicator variables is violated. But also the corrected LISREL-ML standard errors proposed by Jöreskog and Yang (1997) should not be interpreted uncritically for exact inferential statistics or confidence intervals, because in the simulation study they often underestimated the true standard errors of the method, especially when sample size is small or medium.

For the two-stage least squares (2SLS) approach proposed by Bollen (1995, 1996), only the parameters of the structural equation were estimated. The simulation study showed a relatively low bias for standard error estimates, which allowed inferential statistics with acceptable Type I error. Still, the disadvantage of 2SLS lay in its low efficiency relative to the other methods examined. In the study, the application of 2SLS required large sample size in order to provide efficient estimates for the detection of the interaction effect.

The estimation results given in Table 1 are representative for the large study of Schermelleh-Engel, Klein and Moosbrugger (1998) with regard to bias and efficiency of the LMS estimators. Summarizing the results of that study, it can be stated that the LMS parameter estimators indicated to be the most efficient in relation to 2SLS and the LISREL approaches.

5.2. *Model Difference Test*

With a model difference test based on the likelihood ratio test statistic, the interaction model can be tested against a linear structural equation model. The test statistic asymptotically follows a χ^2-distribution, but the distribution of the test statistic may differ for finite sample size. A simulation study was carried out to examine this problem. 800 data sets of sample size $N = 400$ were simulated for the elementary interaction model, with true values taken from Table 1. The linear model is given by restricting the interaction parameter ω_{12} (see (4)) to zero. Every data set was analyzed with LMS, providing a sample of 800 values for the test statistic. Pearson's goodness-of-fit test was carried out to test if this sample followed a χ^2-distribution with d.f. $= 1$. For the goodness-of-fit test, thresholds for the theoretical χ^2-distribution were calculated in order to form 20 categories of equal probability. Then frequencies of the 20 categories were computed from the sample. The result of Pearson's goodness-of-fit test statistic showed no significant deviation of the sample statistics from the theoretical χ^2-distribution ($p = 0.42$). For the model and sample size checked in this study, the model difference test allowed reliable testing of the interaction hypothesis.

If there are no interaction effects, the distribution of the indicator variables is multivariate normal under the assumptions of LMS. Then the models under the null and the alternative hypotheses of the model difference test have different types of distributions which may affect the χ^2-distribution of the test statistic. As the result of the simulation study informed, this aspect had no impact on the distribution of the likelihood ratio test statistic for the elementary interaction model used in the simulation study with sample size $N = 400$.

5.3. *Robustness*

The LMS method assumes the x-variables to be normally distributed, but what happens if these assumptions are violated? To check the robustnes of the method, the following simulation study of the elementary interaction model with nonnormal variables was conducted. The true values of the model parameters used for data generation were taken from Table 1, except the value for the interaction parameter ω_{12}. Four interaction models with four different values of $\omega_{12}(\omega_{12} = 0.0, 0.1, 0.2, 0.7)$ were formed. In the first step, 200 data sets of sample size $N = 400$ for the 7 independent variables $(\xi_1, \xi_2, \delta_1, \delta_2, \delta_3, \delta_4, \zeta)$ were generated with the EQS program (Bentler, 1995). The data for the latent exogenous variable ξ_1 were generated with a skewness of -2.0 and a kurtosis of 6.0. The data for the latent exogenous variable ξ_2 were generated with a skewness of $+1.5$ and a kurtosis of 5.0. The 5 independent error variables $(\delta_1, \delta_2, \delta_3, \delta_4, \zeta)$ were simulated as normally distributed variables. In the second step, the 200 data sets of the 5 indicator variables (x_1, x_2, x_3, x_4, y) were computed for each of the four models separately, according to their model equations. The measurement models for the x-variables are identical in all four models. The skewness and kurtosis of the x-variables depend on their reliabilty. The skewness of x_1, x_2, x_3, x_4 is -0.69, -0.21, 0.77, 0.35, respectively. The kurtosis of x_1, x_2, x_3, x_4 is 1.44, 0.29, 2.05, 0.72, respectively.

TABLE 2.
Percentage of significant test results of the LMS model difference test. The values are based on a simulation study with four elementary interaction models and nonnormal indicator variables. For each of the four models, 200 data sets of sample size $N = 400$ were analyzed. For the Hermite-Gaussian quadrature formula (26) $M = 14$ was chosen.

	LMS Significant model difference test results (in %)
Model with $\omega_{12} = 0.0$	10.4% (Type I Error)
Model with $\omega_{12} = 0.1$	48.0% (Power)
Model with $\omega_{12} = 0.2$	93.5% (Power)
Model with $\omega_{12} = 0.7$	100.0% (Power)

The 200 data sets for each model were analyzed with LMS. A model difference test, which compares the interaction model to the linear model (with ω_{12} set to zero) was executed in LMS for every data set. The theoretical Type I error level of the model difference test was set to 5%. Table 2 reports the percentage of significant test results for the four models.

As can be seen in Table 2, in case of the linear model with interaction parameter zero ($\omega_{12} = 0.0$) LMS provided a significant model difference test in 10.4% of the data sets. Because of the nonnormality of the simulated x-variables, the Type I error level of the model difference test in the study here exceeds the theoretical 5% Type I error level. But, for the degree of nonnormality chosen in this study, there is no dramatic breakdown of the model difference test.

In the model with $\omega_{12} = 0.1$, the interaction effect was detected in only 48% of the cases, because the size of the interaction effect is still relatively low. In the simulated data of the model with $\omega_{12} = 0.2$, the model difference test already provided a significant result in 93.5% of the data sets, although the variance of the interaction term $\omega_{12}\xi_1\xi_2$ determined only 3.8% of the variance of simulated indicator y. Thus, this study gives some evidence that the LMS model difference test is still powerful in case of violated distributional assumptions.

The robustness of LMS with respect to bias end efficiency of the estimates was checked for all four interaction models. It could be expected that the estimation for a model with a relatively high interaction effect size might become critical. The estimation results for the first three models (with $\omega_{12} = 0.0, 0.1, 0.2$) showed no substantial bias in the parameter and standard error estimates. The estimation results for the model with the highest interaction effect size ($\omega_{12} = 0.7$) are reported in Table 3 for the model parameters of the structural equation ($\alpha, \gamma_1, \gamma_2, \omega_{12}$). This table gives for every model parameter: the mean of the parameter estimates (M), the standard deviation of parameter estimates (MC-SD), and the mean of estimated standard errors (Est-SE)

TABLE 3.
Estimation results of a Monte-Carlo study for the elementary interaction model with one latent interaction effect (4). 200 data sets of sample size $N = 400$ were analyzed with LMS and 2SLS. The columns give for the listed model parameters: the true value, the mean of the parameter estimates (M), the standard deviation of parameter estimates (MC-SD), and the mean of estimated standard errors (Est-SE) over all 200 data sets. For the Hermite-Gaussian quadrature formula (26) $M = 14$ was chosen.

		LMS			2SLS		
Parameter	True Value	M	MC-SD	Est-SE	M	MC-SD	Est-SE
α	1.00	0.985	0.041	0.034	1.036	0.112	0.101
γ_1	0.20	0.067	0.100	0.077	0.073	0.299	0.309
γ_2	0.40	0.394	0.069	0.061	0.483	0.175	0.189
ω_{12}	0.70	0.729	0.154	0.110	0.554	0.446	0.380

over all 200 data sets. The 200 data sets for this model were also analyzed with 2SLS (Bollen, 1995, 1996), which is an estimation method that does not assume normally distributed indicator variables.

Whereas the 2SLS estimates are biased for γ_1, γ_1, and ω_{12}, the LMS estimates show a substantial bias only for γ_1. When the estimated standard errors (Est-SE) are compared to the standard deviations computed from the estimates (MC-SD), the study informs that both methods underestimate the MC-SD for ω_{12}. Relative to the size of MC-SD, the estimation of the standard errors by Est-SE provides smaller bias for 2SLS than for LMS. But absolutely, the standard deviations of the parameter estimates (MC-SD) are between two and three times higher for 2SLS compared to LMS: With a mean (M) of 0.554 and a MC-SD of 0.446 for the 2SLS estimates of the interaction parameter ω_{12}, compared to a mean (M) of 0.729 and a MC-SD of 0.154 for LMS in this study, 2SLS did not prove to be a powerful method for the detection of the interaction effect, whereas LMS clearly provided a more efficient parameter estimation. Future simulation research with different models, varied sample size and degree of nonnormality in the variables will decide if these primary results about the robustness of LMS can be generalized.

6. Empirical Example

This section covers an empirical example of LMS data analysis for an interaction model with one latent interaction effect and six indicator variables ($x_1, x_2, x_3, x_4, y_1, y_2$). We used data from Thiele (1998) who investigated age-related effects of coping strategies and the maintaining of well-being for middle-aged males. Part of Thiele's studies concentrated on the effect of flexibility in goal adjustment (ξ_1) on the level of complaining about one's mental or physical situation (η). He formulated an interaction hypothesis suggesting that the level of subjectively perceived fitness (ξ_2) moderates this effect ($\omega_{12} \neq 0$). For persons with a high level of subjective fitness the flexibility of goal adjustment is supposed to have only a small or negligible effect on complaint level, whereas for persons with a low perceived availability of bodily resources, the flexibility of goal adjustment is expected to be an important factor for the level of complaining. Besides this interaction effect, the subjectively perceived fitness is assumed to have a linear additive effect on complaint level. The three variables are latent variables, and their relationship can be modeled by implementing (4):

$$\begin{aligned}\eta &= \alpha + (\gamma_1\ \gamma_2)\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix} + (\xi_1\quad \xi_2)\begin{pmatrix}0 & \omega_{12}\\ 0 & 0\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix} + \zeta, \\ &= \alpha + (\gamma_1 + \omega_{12}\xi_2)\xi_1 + \gamma_2\xi_2 + \zeta.\end{aligned} \tag{34}$$

The structural equation is rewritten in a form which shows that the total effect of ξ_1 on η depends on a linear moderating function ($\gamma_1 + \omega_{12}\xi_2$). The interpretation of γ_1 is *only* reasonable as part of the moderating function (see Moosbrugger, Schermelleh-Engel & Klein, 1997, for a discussion of variable transformation problems in interaction models), which represents the variability of the total effect of ξ_1 on η, moderated by the level of ξ_2.

The flexibility in goal adjustment (ξ_1) refers to the coping style of a person and is indicated by a readiness to disengage from barren commitments, to adapt aspiration levels to feasible range and to find positive meaning in aversive events (accommodative mode of coping). The latent exogenous variable ξ_1 is measured by splitting a flexibility scale from Brandstädter and Renner (1990) into two subscales (x_1, x_2). The perceived fitness (ξ_2) refers to the self-evaluation of the effectiveness with which one's body is functioning. It is also measured as a latent exogenous variable by splitting a scale of self-concept of bodily efficiency (Deusinger, 1998) into two subscales (x_3, x_4). The complaint level is measured by two indicators (psychological complaints y_1, psychovegetative complaints y_2) given by the complaint inventory of Degenhardt and Schmidt (1994), where y_1 measures mental exhaustion and y_2 indicates psychovegetative complaints. The indicators x_1, x_3, and y_1 are used as scaling variables for ξ_1, ξ_2, and η, respectively, with factor loadings set to one. Then the measurement model is given by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda_{x_{21}} & 0 \\ 0 & 1 \\ 0 & \lambda_{x_{42}} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{pmatrix}, \tag{35}$$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda_{y_{21}} \end{pmatrix} \eta + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}. \tag{36}$$

For this elementary interaction model, a data set of sample size $N = 304$ for the joint indicator vector $(\mathbf{x}, \mathbf{y}) = (x_1, x_2, x_3, x_4, y_1, y_2)$ was examined. The data were given in mean deviation form. The univariate skewness of the indicators x_1, x_2, x_3, x_4 was -0.41, -0.34, 0.02, -0.07, respectively. Furthermore, Mardia's coefficient for multivariate kurtosis of the indicator vector $\mathbf{x}$ (Mardia, 1970, 1974) was 0.55 with a critical ratio (kurtosis divided by standard error) of 0.693. So the deviation of $\mathbf{x}$ from normality is not too high, and LMS could be assumed to be robust against this degree of nonnormality (see section 5). The indicator vectors y_1 and y_2 were clearly nonnormal with univariate skewness of 1.23 and 0.92, respectively; their univariate kurtosis was 1.68 and 1.17, respectively. For the joint indicator vector $(\mathbf{x}, \mathbf{y})$, Mardia's coefficient for multivariate kurtosis was 4.43 with a critical ratio of 3.94, which indicates a substantial deviation from normality.

The data of the indicator variables were transformed into standardized scores (with zero mean and standard deviation one) and analyzed by LMS. Table 4 gives the LMS parameter estimates, the estimated standard errors, and the transformed parameter estimates for a completely standardized model.

TABLE 4.
Parameter estimates, estimated standard errors, and parameter estimates for a completely standardized model provided by an LMS analysis of the elementary interaction model with six indicator variables and $N = 304$. For the Hermite-Gaussian quadrature formula (26) $M = 14$ was chosen.

Parameter	Parameter Estimate	Estimated Standard Error	Parameter Estimate for Completely Standardized Model
α	−0.036	0.045	−0.036
γ_1	−0.215	0.061	−0.258
γ_2	−0.457	0.067	−0.493
ω_{12}	0.176	0.059	0.189
Ψ_{11}	0.411	0.061	0.598
$\lambda_{x_{11}}$	1.000	—	0.994
$\lambda_{x_{21}}$	0.601	0.112	0.597
$\lambda_{x_{32}}$	1.000	—	0.894
$\lambda_{x_{42}}$	0.893	0.087	0.798
$\lambda_{y_{11}}$	1.000	—	0.829
$\lambda_{y_{21}}$	1.037	0.088	0.859
ϕ_{11}	0.988	0.179	1.000
ϕ_{21}	0.210	0.055	0.236
ϕ_{22}	0.799	0.105	1.000
$\theta_{\delta_{11}}$	0.009	0.167	0.009
$\theta_{\delta_{22}}$	0.640	0.080	0.640
$\theta_{\delta_{33}}$	0.208	0.070	0.208
$\theta_{\delta_{44}}$	0.367	0.061	0.367
$\theta_{\varepsilon_{11}}$	0.302	0.056	0.302
$\theta_{\varepsilon_{22}}$	0.250	0.057	0.250

The negative signs of γ_1 and γ_2 confirm the expected linear effects that the complaint level is low when goal adjustment is flexible or subjectively perceived fitness is high. The *unstandardized* interaction parameter estimate $\hat{\omega} = 0.176$ with standard error $\mathrm{SE}(\hat{\omega}_{12}) = 0.059$ yields the confidence interval [0.06, 0.29] for the 5% Type I error level. Therefore, the interaction parameter is significantly different from zero. The investigation of the *standardized* moderating function $(\hat{\gamma}_1 + \hat{\omega}_{12}\xi_2) = (-0.258 + 0.189\xi_2)$ shows that the latent interaction neutralizes the effect of flexibility of goal adjustment ξ_1 on η if the subjectively perceived fitness ξ_2 has a high level, whereas for a low level of fitness the flexibility level of goal adjustment has a substantial impact on complaint level.

Parallel to the LMS analysis, the empirical data set was analyzed with LISREL-ML. LISREL-ML provided the interaction parameter estimate $\hat{\omega}_{12} = 0.19$ with standard error $\mathrm{SE}(\hat{\omega}_{12}) = 0.14$. This yields the confidence interval $[-0.08, 0.46]$ for the 5% Type I error level. Therefore, the estimate calculated by LISREL-ML is not significantly different from zero and the interaction effect cannot be detected in this case.

Finally, a model difference test based on the likelihood-ratio test statistic for ML estimation confirms that the interaction model (model hypothesis H_1) fits the data significantly better than a linear model with interaction parameter ω_{12} set to zero (model hypothesis H_0). The chi-square value of the model difference test was $\chi^2_{\mathrm{diff,df}=1} = 9.7$ with p-value less than 0.01. In this empirical example, the high efficiency of LMS estimators and the unbiased estimation of their standard errors provide a reliable device for the detection of latent interaction effects.

7. Conclusion

The LMS estimation procedure takes the distributional characteristics of the nonnormally distributed joint indicator vector in a latent interaction model explicitly into account. By analyzing the density function, the LMS approach implements an iterative ML estimation procedure tailored for the type of nonnormality induced by interaction effects and additionally gives theoretical insight into the stochastic structure of latent interaction. Our simulation study suggests that the LMS parameter estimation is unbiased and more efficient than the estimation with alternative estimation techniques. Just as important, the estimation of standard errors in LMS showed no substantial bias which supports precise hypothesis testing of interaction effects. A model difference test for LMS testing the interaction model against a linear model is obtained by applying the general likelihood ratio test statistic for ML estimators. A first simulation study indicates that the method is robust against moderate violation of its distributional assumptions. LMS seems to be a theoretically promising and practically adequate approach to the analysis of latent interaction effects.

References

Abramowitz, M., & Stegun, I. A. (1971). *Handbook of mathematical functions*. New York, NY: Dover Publications.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park: SAGE Publications.

Arbuckle, J. L. (1997). *AMOS Users' Guide Version 3.6*. Chicago: Small Waters Corporation.

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

Bentler, P. M., & Wu, E. J. C. (1993). *EQS/Windows user's guide*. Los Angeles: BMDP Statistical Software.

Bollen, K. A. (1995). Structural equation models that are nonlinear in latent variables. In P. V. Marsden (Ed.), *Sociological methodology 1995* (Vol. 25). Washington, DC: American Sociological Association.

Bollen, K. A. (1996). An alternative two stage least squares (2SLS) estimator for latent variable equations. *Psychometrika, 61*, 109–121.

Brandstädter, J., & Renner, G. (1990). Tenacious goal pursuit and flexible goal adjustment: Explication and age-related analysis of assimilative and accomodative strategies of coping. *Psychology and Aging, 5*, 58–67.

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analyses for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Degenhardt, A., & Schmidt, H. (1994). Physische Leistungsvariablen als Indikatoren für die Diagnose "Klimakterium Virile" (Physical efficiency variables as indicators for the diagnosis of 'climacterium virile'). *Sexuologie, 3*, 131–141.

Dempster, A.P., Laird, N.M., & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39*, 1–38.

Deusinger, I. (1998). *Frankfurter Körperkonzept-Skalen* [Frankfurt bodily self-concept scales]. Göttingen: Hogrefe.

Grim, J. (1982). On numerical evaluation of maximum-likelihood estimates for finite mixtures of distributions. *Kybernetika, 18*(3), 173–190.

Hayduk, L.A. (1987). *Structural equation modeling with LISREL*. Baltimore, MD: Johns Hopkins University Press.

Isaacson, E., & Keller, H.B. (1966). *Analysis of numerical methods*. New York, NY: Wiley.

Jaccard, J., Turrisi, R., & Wan, C.K. (1990). *Interaction effects in multiple regression*. Newbury Park, CA: Sage Publications.

Jöreskog, K.G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications* (2nd ed.). Chicago, IL: SPSS.

Jöreskog, K.G., & Sörbom, D. (1993). *New features in LISREL 8*. Chicago, IL: Scientific Software.

Jöreskog, K. G., & Sörbom, D. (1996). *PRELIS 2: User's guide*. Chicago: Scientific Software.

Jöreskog, K.G., & Yang, F. (1996). Nonlinear structural equation models: The Kenny-Judd model with interaction effects. In G. Markoulides & R. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 57–87). Mahwah, NJ: Lawrence Erlbaum Associates.

Jöreskog, K. G., & Yang, F. (1997). Estimation of interaction models using the augmented moment matrix: Comparison of asymptotic standard errors. In W. Bandilla & F. Faulbaum (Eds.), *SoftStat '97. Advances in statistical software 6* (pp. 467–478). Stuttgart: Lucius & Lucius.

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin, 96*, 201–210.

Klein, A., Moosbrugger, H., Schermelleh-Engel, K., & Frank, D. (1997). A new approach to the estimation of latent interaction effects in structural equation models. In W. Bandilla & F. Faulbaum (Eds.), *SoftStat '97. Advances in statistical software 6* (pp. 479–486). Stuttgart: Lucius & Lucius.

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 in testing normality and robustness studies. *Sankhya, Series B, 36*, 115–128.

Moosbrugger, H., Frank, D., & Schermelleh-Engel, K. (1991). Zur Überprüfung von latenten Moderatoreffekten mit linearen Strukturgleichungsmodellen [Estimating latent interaction effects in structural equation models]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 12*, 245–255.

Moosbrugger, H., Schermelleh-Engel, K., & Klein, A. (1997). Methodological problems of estimating latent interaction effects. *Methods of Psychological Research Online, 2*, 95–111.

Ping, R. A. (1996a). Latent variable and quadratic effect estimation: A two-step technique using structural equation analysis. *Psychological Bulletin, 119*, 166–175.

Ping, R. A. (1996b). Latent variable regression: A technique for estimating interaction and quadratic coefficients. *Multivariate Behavioral Research, 31*, 95–120.

Redner, R. A., & Walker, H. F. (1984). Mixture densities, maximum likelihood and the EM algorithm. *SIAM Review, 26*, 195–239.

Schermelleh-Engel, K., Klein, A., & Moosbrugger, H. (1998). Estimating nonlinear effects using a Latent Moderated Structural Equations Approach. In R. E. Schumacker & G. A. Marcoulides (Eds.), *Interaction and nonlinear effects in structural equation modeling* (pp. 203–238). Mahwah, NJ: Lawrence Erlbaum Associates.

Schmitt, M. (1990). *Konsistenz als Persönlichkeitseigenschaft? Moderatorvariablen in der Persönlichkeits- und Einstellungsforschung* [Consistency as a personality trait? Moderator variables in personality and attitude research]. Berlin: Springer.

Schwarz, H. R. (1993). *Numerische Mathematik* [Numerical mathematics]. Stuttgart: Teubner.

Thiele, A. (1998). *Verlust körperlicher Leistungsfähigkeit: Bewältigung des Alterns bei Männern im mittleren Lebensalter* [Loss of bodily efficacy: The coping of aging for men of medium age]. Idstein, Germany: Schulz-Kirchner-Verlag.

Yang Jonsson, F. (1997). *Nonlinear structural equation models: Simulation studies of the Kenny-Judd model*. Uppsala: University of Uppsala.

Manuscript received 23 MAR 1998
Final version received 30 MAR 1999