Review of Psychology,
2009, Vol. 16, No. 2, 123-131 UDC 159.9

Estimation of nonlinear latent structural equation models using the extended unconstrained approach

AUGUSTIN KELAVA and HOLGER BRANDT

In the past two decades, several approaches for the estimation of single interaction effects in latent structural equation models (SEM) have been published (for an overview: Marsh, Wen, & Hau, 2006; Schumacker & Marcoulides, 1998). Recent literature has begun to consider nonlinear models involving simultaneous interaction and quadratic effects (e.g. Klein & Muthén, 2007; Lee, Song, & Poon, 2004; MacCallum & Mar, 1995). Marsh, Wen, and Hau (2004) developed an unconstrained product indicator approach for the estimation of single interaction effects with robust properties when distributional assumptions are violated. In this article, first, we give a brief overview on the approaches for the estimation of nonlinear SEM. Second, we describe the extended unconstrained approach for the simultaneous estimation of latent interaction and quadratic effects (Kelava, 2009; Moosbrugger, Schermelleh-Engel, Kelava, & Klein, 2009). Third, we apply the extended unconstrained approach to data from work and stress research using the freely accessible sem package (Fox, 2006) in R (R Development Core Team, 2008) and compare it with LMS (Klein & Moosbrugger, 2000) which is implemented in the commercial M*plus* (Muthén & Muthén, 2007) software.

Key words: nonlinear structural equation models, estimators, quadratic, interaction, unconstrained approach

Numerous theories within the social and behavioral sciences hypothesize interaction, quadratic effects, or both between multiple independent and dependent variables (Ajzen, 1987; Cronbach & Snow, 1977; Karasek, 1979; Lusch & Brown, 1996; Snyder & Tanke, 1976). For example, Ganzach (1997) studies the relationship between parents' educational level and child's educational expectations. He hypothesizes and finds a simultaneous interactive and quadratic relationship: If at least one parent's education level is high, the educational expectations of the child will also be high, even if the level of education of the other parent is quite low. In terms of the statistical model, this compensatory hypothesis is represented by two positive quadratic effects (for each parent's educational level) and one negative interaction effect. Within the measured variable framework, such hypotheses can be tested by specifying a multiple regression model (see Aiken & West, 1991):

$$CEE = \beta_0 + \beta_1 ME + \beta_2 FE + \omega_{12} ME \bullet FE + \omega_{11} ME^2 + \omega_{22} FE^2 + \varepsilon \quad (1)$$

where *CEE* is the child's educational expectation, *ME* is the mother's educational level, *FE* is the father's educational level, and ε is a residual. The γs are the coefficients of the linear effects. Following Klein and Moosbrugger's (2000) and Klein and Muthén's (2007) notation, the ωs are the coefficients of the nonlinear effects.

To clarify the necessity for models with simultaneous interaction and quadratic effects consider, for example, Ganzach's compensatory hypothesis. The hypothesis that only one parent's educational level needs to be high for high educational expectations of the child could not be tested with an (ordinary) single interaction effect model. A model with a single interaction effect would predict that each parent's education has to be high for a high educational expectation of the child. This would be the interpretation of a positive interaction effect which would result if the quadratic terms were omitted in the analysis[1]. In Equation (1), the opposite

Augustin Kelava, Department of Psychology, Technical University Darmstadt, Darmstadt, Institute of Psychology, S1 15, Alexanderstrasse 10, D-64283 Darmstadt, Germany. E-Mail: tino@augustin-kelava.de (the address for correspondence);

Holger Brandt, Department of Psychology, Goethe University Frankfurt, Frankfurt am Main, Germany.

1 The true negative interaction effect and the two positive quadratic effects could reduce to one single positive interaction effect when the quadratic effects are omitted.

signs of the quadratic and interaction effects assure the compensatory relationship.

Since most variables in the behavioral sciences are measured with less than perfect reliability, a regression analysis often is not appropriate. Having non-perfectly reliable predictors results in biased estimates of the regression coefficients, especially for the nonlinear effects (Bohrnstedt & Marwell, 1978; MacCallum & Mar, 1995). Structural equation modeling (SEM) produces theoretically error free estimates of the effects of latent variables, overcoming this reliability problem (Marsh et al., 2004; Schumacker & Marcoulides, 1998). But, structural equation modeling of interaction and quadratic effects has rarely been used by practitioners. This is partly due to the error-prone model specification within the traditional product indicator approaches (for an overview: Marsh et al., 2006). And, it is partly due to the necessity to use specialized commercial software, e.g. LISREL (Jöreskog & Sörbom, 1996) or M*plus* (Muthén & Muthén, 2007), in order to specify and estimate nonlinear SEM.

Goals of the article

The major goals of this article are threefold: First, we will give a brief overview on the approaches for the estimation of nonlinear SEM. Second, we describe the extended unconstrained approach for the simultaneous estimation of latent interaction and quadratic effects (Kelava, 2009; Moosbrugger, Schermelleh-Engel, Kelava, & Klein, 2009). Third, we will apply the extended unconstrained approach to data from work and stress research using the freely accessible sem package (Fox, 2006) in R (R Development Core Team, 2008) and compare it with LMS (Klein & Moosbrugger, 2000) which is implemented in the commercial Mplus (Muthén & Muthén, 2007) software. Example syntax will be given in the Appendices.

Approaches for the estimation of nonlinear SEM

Most of the early literature focused on models with a single latent interaction or quadratic effect (e.g. Jöreskog & Yang, 1996; Kenny & Judd, 1984). Recently, the literature has begun to consider more complex models like Ganzach's (1997) model of children's educational expectations involving simultaneous interaction and quadratic effects (Kelava, Moosbrugger, Dimitruk, & Schermelleh-Engel, 2008; Kelava et al., under revision; Klein & Muthén, 2007; Lee et al., 2004; Lee, Song, & Tang, 2007; MacCallum & Mar, 1995). Equation (2) expresses Ganzach's model with one interaction and two quadratic effects (see Equation (1)) within the latent variable framework:

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1 \cdot \xi_2 + \omega_{11}\xi_1^2 + \omega_{22}\xi_2^2 + \zeta \quad (2)$$

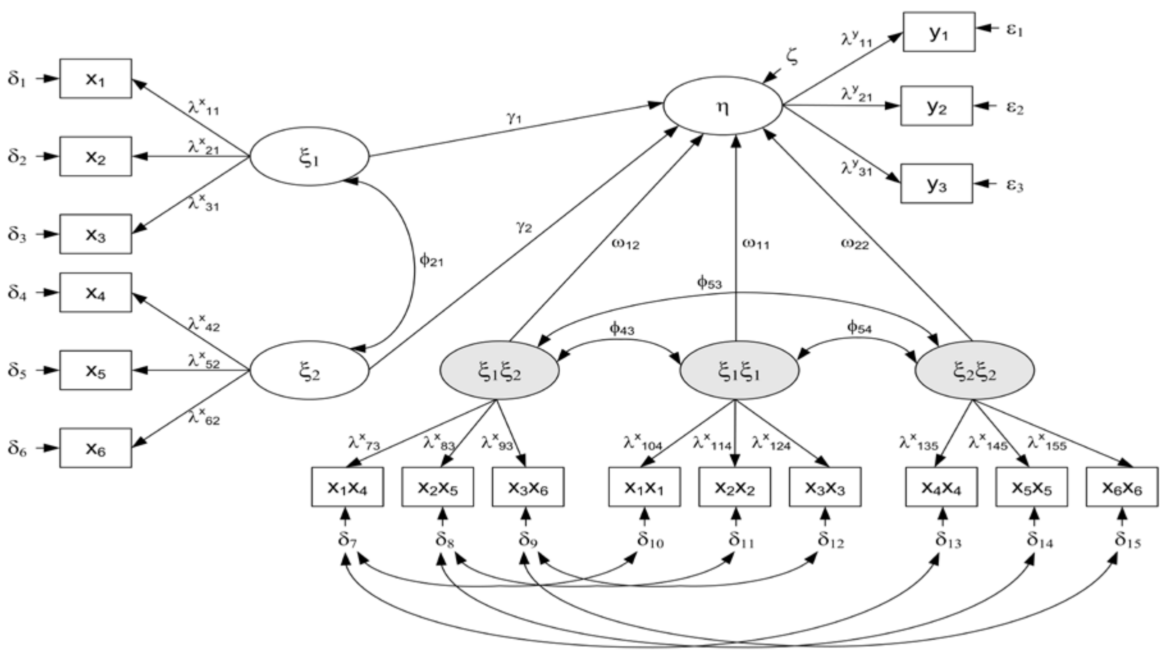

Figure 1. Nonlinear SEM with one interaction effect and two quadratic effects. Each latent variable is operationalized by three indicators. Within the nonlinear measurement model, the measurement error covariances have to be specified, when the latent linear predictors ξ_1 and ξ_2 are correlated (Kelava, 2009).

In Equation (2), η denotes the latent criterion, ξ_1 and ξ_2 are latent predictors, the product $\xi_1\xi_2$ represents the interaction term, ξ^2_1 and ξ^2_2 are quadratic terms, α is the intercept, γ_1 and γ_2 are linear effects of the predictors, ω_{12} is the nonlinear effect of the interaction term, and ω_{11}, ω_{22} are the nonlinear effects of the quadratic terms, and finally, ζ is the disturbance term. Figure 1 depicts this nonlinear SEM model with one interaction effect and two quadratic effects.

Kenny and Judd (1984) were the first who developed an approach for the estimation of nonlinear SEM. It is called product indicator approach, because multiple product indicators are used for the specification of each nonlinear term's measurement model. Suppose that the normally distributed and centered latent variables ξ_1 and ξ_2 are measured by centered indicators x_1, x_2, x_3 and x_4, x_5, x_6, respectively (Equation (3)):

$$\mathbf{x} = \mathbf{\Lambda^x} \cdot \xi + \delta$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda^x_{21} & 0 \\ \lambda^x_{31} & 0 \\ 0 & 1 \\ 0 & \lambda^x_{52} \\ 0 & \lambda^x_{62} \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{pmatrix} \tag{3}$$

whereas the λ^x s are factor loadings and the δ s are normally distributed measurement errors. The interaction term $\xi_1\xi_2$ is measured by products of each latent variable's indicators, for example x_1x_4, x_2x_5, x_3x_6 (see Figure 1). The quadratic term ξ^2_1 is measured by x_1^2, x_2^2, x_3^2, and so on. Unfortunately, this approach (in its original form) has rarely been used by applied researchers. The main reason is that it involves the specification of nonlinear parameter constraints that are difficult for researchers to implement. Suppose that x_2 and x_5 are indicators of the linear latent predictor variables ξ_1 and ξ_2 (with $x_2 = \lambda^x_{21}\xi_1 + \delta_2$ and $x_5 = \lambda^x_{52}\xi_2 + \delta_5$), then the indicator x_2x_5 of the interaction term $\xi_1\xi_2$ would be:

$$\begin{aligned} x_2x_5 &= \underbrace{\lambda^x_{21}\lambda^x_{52}}_{\lambda^x_{83}}\xi_1\xi_2 + \underbrace{\lambda^x_{52}\xi_2\delta_2 + \lambda^x_{21}\xi_1\delta_5 + \delta_2\delta_5}_{\delta_8} \\ &= \lambda^x_{83}\xi_1\xi_2 + \delta_8 \end{aligned} \tag{4}$$

The variance decomposition of the indicator product x_2x_5 which is required for the model specification and estimation (for example) in the LISREL software, is given by:

$$Var(x_2x_5) = \lambda^{x\,2}_{83} Var(\xi_1\xi_2) + Var(\delta_8) \tag{5}$$

where:

$$\lambda^x_{83} = \lambda^x_{21}\lambda^x_{52} \tag{6}$$

$$Var(\xi_1\xi_2) = Var(\xi_1)Var(\xi_2) + Cov^2(\xi_1, \xi_2) \tag{7}$$

$$\begin{aligned} Var(\delta_8) = {} & \lambda^{x\,2}_{21} Var(\xi_1)Var(\delta_5) + \lambda^{x\,2}_{52} Var(\xi_2)Var(\delta_2) \\ & + Var(\delta_2)Var(\delta_5) \end{aligned} \tag{8}$$

Because factor loadings and variances of the indicator products are functions of the factor loadings and variances of the linear indicators, this estimation approach demands the specification of nonlinear parameter constraints, which is very error prone. Furthermore these constraints only hold if the latent predictors are normally distributed (see Wall & Amemiya, 2001).

Fortunately, two different trends recently emerged, one trying to simplify and expand the product indicator approach, which is the biggest class of approaches, and one coming from a different perspective and using a distribution-analytic approach.

The product indicator approach was particularly developed as reflected in contributions by Jaccard and Wan (1995), Ping (1995, 1996), Jöreskog and Yang (1996), Algina and Moulder (2001), Wall and Amemiya (2001), Marsh et al. (2004), Little, Bovaird, and Widaman (2006). These developments led to simplifications of the specified model. The simplest approach has been published by Marsh et al. (2004). The "unconstrained approach" was developed for the estimation of single interaction effects. It relaxes all nonlinear constraints. This means that, for example, $Var(\delta_8)$ in Equation (8) is not constrained to the right-hand side combination of the parameters, but instead is estimated freely. In the next section, we will go into detail about this approach and about its extension for the simultaneous estimation of interaction and quadratic effects.

Since the traditional product indicator approach suffers from the violated assumption of multivariate normally distributed variables[2] when ML estimates are derived, so-called distribution-analytic approaches have been developed that address the nonnormal distribution. Klein and Moosbrugger (2000) developed a Latent Moderated Structural Equations (LMS) approach which approximates the nonnormal distribution of the multivariate indicator vector by a mixture of normal distributions. By applying the EM algorithm (Dempster, Laird, & Rubin, 1977), ML estimates are obtained. LMS computes unbiased standard errors for the nonlinear effects, which are slightly underestimated when applying the product indicator approach (Jöreskog & Yang, 1996; Kelava et al., 2008). Unfortunately, this approach becomes computationally (numerically) intensive as the number of

2 When applying the ML estimator, it is assumed that the product indicators (e.g., x_1x_4) and the y-indicators are normally distributed. This assumption never holds, because products of normally distributed variables are never normally distributed (Aroian, 1944).

nonlinear effects increases. In addition to this, LMS is only available within the commercial Mplus software, but not available within a freely accessible software. In order to overcome the problem of high computational burden and in order to develop a more robust approach when indicators are nonnormally distributed, Klein and Muthén (2007) published a Quasi-Maximum Likelihood (QML) approach. QML permits the estimation of multiple nonlinear effects. It approximates the likelihood of the multivariate nonnormally distributed indicator vector by a normal and a conditionally normal distribution. The parameter estimates are obtained by using a Newton-Raphson algorithm.

In addition to the distribution-analytic approaches (LMS and QML) and product indicator approaches, a variety of less-established, alternative approaches has been developed (for an overview: Marsh et al., 2006; Schumacker & Marcoulides, 1998). For example, there are also Bayesian approaches (Arminger & Muthén, 1998; Lee et al., 2007), a 2-step method of moments (2SMM) approach (Wall & Amemiya, 2000), and a 2-step least squares (2SLS) approach (Bollen, 1995). Elaborated simulation studies need to be conducted with the Bayesian approaches and the 2SMM approach in order to assess the robustness and competitiveness with the established approaches. 2SLS estimates were substantially less efficient when compared to alternative estimation approaches (Klein & Moosbrugger, 2000; Schermelleh-Engel, Klein, & Moosbrugger, 1998).

In the following, we will describe the well-known unconstrained approach (Marsh et al., 2004) and its extension for the simultaneous estimation of interaction and quadratic effects. The unconstrained approach can be implement by using commercial structural equation modeling software (e.g., LISREL) or by using the freely available sem package in R (see Appendix A). The unconstrained approach has proven to have robust properties in specific circumstances (Kelava, 2009; Kelava et al., under revision; Marsh et al., 2004, 2006).

The extended unconstrained approach

In this section, we will summarize the unconstrained approach for the estimation of single interaction effects as proposed by Marsh et al. (2004, 2006). After this we will present an extension of the unconstrained approach for the simultaneous estimation of interaction and quadratic effects as proposed by Kelava (2009) and Moosbrugger et al. (2009). We provide a detailed Technical Appendix showing how to estimate the models using the freely accessible sem package (Fox, 2006) in R (R Development Core Team, 2008).

Marsh et al.'s (2004) unconstrained approach

The main idea of the unconstrained approach is to relax all constraints formulated by Kenny and Judd (1984) that make the specification of the interaction model complicated (see Equations (5) - (8)) and to estimate these parameters freely. Suppose we have a simple interaction model:

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \zeta \tag{9}$$

with normally ξ_1, ξ_2 distributed and ζ variables with means equal to zero. The linear measurement models are given by:

$$\mathbf{x} = \mathbf{\Lambda}^{\mathbf{x}} \cdot \xi + \delta$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda^x_{21} & 0 \\ \lambda^x_{31} & 0 \\ 0 & 1 \\ 0 & \lambda^x_{52} \\ 0 & \lambda^x_{62} \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{pmatrix} \tag{10}$$

and

$$\mathbf{y} = \mathbf{\Lambda}^{\mathbf{y}} \cdot \eta + \varepsilon$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda^y_{21} \\ \lambda^y_{31} \end{pmatrix} \cdot \eta + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} \tag{11}$$

with centered and normally distributed δ and ε measurement error variables. The nonlinear measurement model is given by:

$$\begin{pmatrix} x_1x_4 \\ x_2x_5 \\ x_3x_6 \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda^x_{83} \\ \lambda^x_{93} \end{pmatrix} \cdot \xi_1\xi_2 + \begin{pmatrix} \delta_7 \\ \delta_8 \\ \delta_9 \end{pmatrix} \tag{12}$$

Although we assume that $\kappa_1 := E(\xi_1) = 0$ and $\kappa_2 := E(\xi_2) = 0$ the latent expectation $\kappa_3 = E(\xi_1\xi_2)$ will be equal to $Cov(\xi_1\xi_2)$ and thus needs to be estimated (or otherwise constrained).

In summary, since parameter estimation is based on empirical and model implied covariance matrices (like in LISREL or in the sem package), the following parameters have to be estimated freely:

- regression coefficients: $\alpha, \gamma_1, \gamma_2, \omega_{12}, \lambda^x_{21}, \lambda^x_{31}, \lambda^x_{52}, \lambda^x_{62}, \lambda^x_{83}, \lambda^x_{93}, \lambda^y_{21}$, and λ^y_{31}
- variances and covariances of the latent predictors: $Var(\xi_1), Cov(\xi_2, \xi_1), Var(\xi_2)$, and $Var(\xi_1\xi_2)$
- variances of the disturbances: $Var(\zeta), Var(\delta_1), Var(\delta_2), ..., Var(\delta_9), Var(\varepsilon_1), Var(\varepsilon_2)$, and $Var(\varepsilon_3)$
- latent expectation of the nonlinear predictor: κ_3 (or otherwise constrained to $Cov(\xi_1, \xi_2)$)

For purposes of model identification, $\lambda^x_{11}, \lambda^x_{42}, \lambda^x_{73}$, and λ^y_{11} have to be fixed at 1 and need not to be estimated.

Instead of estimating the structural model's latent intercept α and the latent expectation κ_3, latent intercepts of the outcome indicators (y_1, y_2, y_3) and product indicators (x_1x_4, x_2x_5, x_3x_6) can be estimated, too. For example, α could be omitted in Equation (9). Then, the latent intercepts τ^y_1, τ^y_2, and τ^y_3 must be estimated within a modified outcome measurement model:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \tau^y_1 \\ \tau^y_2 \\ \tau^y_3 \end{pmatrix} + \begin{pmatrix} 1 \\ \lambda^y_{21} \\ \lambda^y_{31} \end{pmatrix} \cdot \eta + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} \tag{13}$$

In the modified product indicator measurement model, then τ^x_1, τ^x_2, and τ^x_3 must be estimated, when omitting κ_3:

$$\begin{pmatrix} x_1x_4 \\ x_2x_5 \\ x_3x_6 \end{pmatrix} = \begin{pmatrix} \tau^x_1 \\ \tau^x_2 \\ \tau^x_3 \end{pmatrix} + \begin{pmatrix} 1 \\ \lambda^x_{83} \\ \lambda^x_{93} \end{pmatrix} \cdot \xi_1\xi_2 + \begin{pmatrix} \delta_7 \\ \delta_8 \\ \delta_9 \end{pmatrix} \tag{14}$$

If the indicator variables $x_1,\ldots,x_6$ are nonnormally distributed, Marsh et al. (2004) propose to estimate $Cov(\xi_1\xi_2,\xi_1)$ and $Cov(\xi_1\xi_2,\xi_2)$ additionally, because these covariances are not equal to zero, if the latent predictors ξ_1 and ξ_2 are nonnormally distributed. If it can be assumed that nonnormality results from nonnormally distributed measurement error variables of the linear indicators (e.g., ceiling effects can produce a nonnormal δ_2 measurement error variable), then measurement error covariances between the linear and their related nonlinear indicators (e.g., $Cov(\delta_2,\delta_8)$) should be specified and estimated, too (because both, x_2 and $x_2x_5 =: x_8$, contain the nonnormal δ_2 which is also part of δ_8). A nonlinear SEM with a single quadratic effect, instead of an interaction effect, is specified analogously.

The extended unconstrained approach for the simultaneous estimation of latent interaction and quadratic effects

In this subsection we present the extension of the unconstrained approach for the simultaneous estimation of interaction and quadratic effects. In order to keep it as simple as possible, *we will assume that the latent predictors ξ_1 and ξ_2 and their indicators are normally distributed and centered (with zero means).*

Equation (15) shows the nonlinear SEM with both effect types:

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1 \cdot \xi_2 + \omega_{11}\xi_1^2 + \omega_{22}\xi_2^2 + \zeta \tag{15}$$

The linear measurement models for ξ_1, ξ_2 and η are given by Equations (10) and (11). The nonlinear measurement model is given by (cp. Figure 1):

$$\begin{pmatrix} x_1x_4 \\ x_2x_5 \\ x_3x_6 \\ x_1^2 \\ x_2^2 \\ x_3^2 \\ x_4^2 \\ x_5^2 \\ x_6^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \lambda^x_{83} & 0 & 0 \\ \lambda^x_{93} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda^x_{114} & 0 \\ 0 & \lambda^x_{124} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda^x_{143} \\ 0 & 0 & \lambda^x_{153} \end{pmatrix} \cdot \begin{pmatrix} \xi_1\xi_2 \\ \xi_1^2 \\ \xi_2^2 \end{pmatrix} + \begin{pmatrix} \delta_7 \\ \delta_8 \\ \delta_9 \\ \delta_{10} \\ \delta_{11} \\ \delta_{12} \\ \delta_{13} \\ \delta_{14} \\ \delta_{15} \end{pmatrix} \tag{16}$$

Once again, $\kappa_3=E(\xi_1\xi_2)$, $\kappa_4=E(\xi^2_1)$ and $\kappa_5=E(\xi^2_2)$ are not equal to zero in general. These expectations of the nonlinear latent variables have to be constrained or estimated freely.

When estimating a nonlinear SEM with simultaneous interaction and quadratic effects, additional measurement error covariances of the nonlinear indicators (e.g. $Cov(\delta_8,\delta_{11})$) must be specified, because they are not zero when $Cov(\xi_1,\xi_2) \neq 0$ (Kelava, 2009; Kelava et al., 2008). This should be the case in most research situations. *If the covariances are not specified, the estimates of the nonlinear effects will be severely biased* (Kelava, 2009). Unfortunately, these additional covariances have been omitted not only in early literature (Kenny & Judd, 1984), but also in recent literature (Lee et al., 2004). Suppose that x_2x_5 is an indicator of $\xi_1\xi_2$ and x_2 is an indicator of ξ^2_1 (see Figure 1). Then $Cov(\delta_8,\delta_{11})$ must be estimated, because x_2's measurement error δ_2 is part of x_2x_5 (see Equation (4)) and part of x_2^2). In Figure 1, we have used double-sided arrows to show that the measurement error covariances have to be estimated freely.

As long as variables are centered and normally distributed, measurement error covariances between linear and nonlinear indicators (e.g., $Cov(\delta_2,\delta_8)$) need not to be estimated, because they are third moments and are equal to zero. When the variables are nonnormally distributed, they will not be equal to zero and need to be estimated, too (assuming the model being identified).

Within the extended unconstrained approach, the following parameters have to be estimated freely:

- regression coefficients: $\alpha, \gamma_1, \gamma_2, \omega_{12}, \omega_{11}, \omega_{22}, \lambda^x_{21}, \lambda^x_{31}, \lambda^x_{52}, \lambda^x_{62}, \lambda^x_{83}, \lambda^x_{93}, \lambda^x_{114}, \lambda^x_{124}, \lambda^x_{145}, \lambda^x_{155}, \lambda^y_{21}$, and λ^y_{31}
- variances and covariances of the latent predictors: $Var(\xi_1), Cov(\xi_2,\xi_1), Var(\xi_2), Var(\xi_1\xi_2), Cov(\xi_1^2,\xi_1\xi_2), Var(\xi_1^2), Cov(\xi_2^2,\xi_1\xi_2), Cov(\xi_2^2,\xi_1^2)$, and $Var(\xi_2^2)$

- variances of the disturbances: $Var(\zeta), Var(\delta_1), Var(\delta_2), ..., Var(\delta_{15}), Var(\varepsilon_1), Var(\varepsilon_2), Var(\varepsilon_3), Cov(\delta_7, \delta_{10}), Cov(\delta_8, \delta_{11}), Cov(\delta_9, \delta_{12}), Cov(\delta_7, \delta_{13}), Cov(\delta_8, \delta_{14})$, and $Cov(\delta_9, \delta_{15})$
- latent expectations of the nonlinear predictors: κ_3, κ_4 and κ_5.

For purposes of model identification $\lambda_{11}^x, \lambda_{42}^x, \lambda_{73}^x, \lambda_{104}^x, \lambda_{135}^x$, and λ_{11}^y have to be fixed at 1 and need not to be estimated.

As before, instead of estimating α and κ_3 - κ_5, latent intercepts τ_1^y, τ_2^y, and τ_3^y of the outcome indicators y_1, y_2, and y_3, and latent intercepts $\tau_1^x, \tau_2^x, .., \tau_9^x$ of the nonlinear product indicators $x_1x_4, ..., x_1^2, x_2^2, ..., x_6^2$ can be specified, too (cp. Equation (13) and (14)).

Empirical example from the work and stress research

In this section, we provide a very brief empirical example from the work and stress research which is based on a data set that has already been published in a larger publication by Diestel and Schmidt (2009)[3]. In their publication, the authors examined the relationship between 'work load' (wl) and 'anxiety' (ax) and found that the variable 'demands for impulse control' (ic) is a significant moderator of that relationship. In detail, they found that higher work load and higher demands to control emotional impulses lead to increased anxiety. In addition to these linear effects, higher work load leads to higher anxiety when the demands to control emotional impulses are high.

In terms of a regression model, this result can be expressed by the following Equation (17):

$$ax = .551wl + .185ic + .155wl \cdot ic + \zeta \qquad (17)$$

where the explained variance is R^2 = .492 and N = 574.

The (standardized) data were analyzed with LMS (Klein & Moosbrugger, 2000) which is implemented in the the commercial M*plus* (Muthén & Muthén, 2007) software. As can be seen from Equation (17), there is a strong effect of work load and a relatively high interaction effect of work load and demands for impulse control.

We reanalyzed the original data set and specified an additional quadratic effect for the demands for impulse control. We applied two approaches: First, we used the extended unconstrained approach which can be implemented in the non-commercial sem package (Fox, 2006) in R (R Development Core Team, 2008). Second, we analyzed the data with the additional effect using the LMS approach. A detailed description on how to apply both approaches is given in Appendix A (sem package) and Appendix B (M*plus*).[4]

3 We gratefully thank Stefan Diestel for sharing the original data.

4 Appendices are available at the journal webpage http://psihologija.ffzg.hr/review.

Table 1

Results of the reanalysis of the Diestel and Schmidt (2009) data with the extended unconstrained approach and the LMS approach

Approach	Parameter	Estimate	Standard error	z-value	p
Extended unconstrained approach	γ_1	.534	.060	8.881	< .001
	γ^2	.171	.043	4.018	< .001
	ω_{12}	.238	.083	2.853	.004
	ω_{22}	-.071	.041	-1.727	.084
LMS approach	γ_1	.545	.061	8.974	< .001
	γ_2	.166	.041	4.014	< .001
	ω_{12}	.190	.055	3.487	< .001
	ω_{22}	-.072	.037	-1.943	.052

In order to illustrate the procedure for both approaches, a hypothetical data set was generated and analyzed.

The analyzed model can be summarized by the following Equation (18):

$$ax = \gamma_1 wl + \gamma_2 ic + \omega_{12} wl \cdot ic + \omega_{22} ic^2 + \zeta \qquad (18)$$

Results are given in Table 1. As can be seen, in both approaches, there are significant linear effects and a significant interaction effect (according to the analyses of Diestel & Schmidt, 2009). But, in LMS there is also a significant quadratic effect of the demands for impulse control. In the unconstrained approach, we were modeling the additional (proposed) measurement error covariances of the product indicators and linear indicators, in order to account for the non-normality in the data, and found that there is no significant quadratic effect. Since Diestel and Schmidt (2009) report substantive non-normality in the data, the significant additional quadratic effect in LMS might be spurious. Simulation studies (Brandt, 2009) and theoretical considerations (Klein & Muthén, 2007; Kelava et al., under revision) have shown that LMS (but not QML!) should be more vulnerable to non-normality (due to its distributional assumptions).

DISCUSSION

In this article three goals were set. First, we gave a short overview on the different types of approaches for the estimation of nonlinear structural equation models. Mainly two approaches have shown to be easily applicable by applied researchers. While the application of the distribution analytic approaches has been described in Kelava et al. (under revision), the user-oriented description on how to apply the product-indicator approach when estimating multiple nonlinear effects has not been published, by now.

Second, we described the unconstrained approach in detail (Marsh et al., 2004) and extended the original model with one interaction effect to a model with one interaction effect and two quadratic effects, because the unconstrained approach has shown to have some robust properties (Kelava, 2009; Marsh et al., 2004).

Third, we provided a brief example from the work and stress research finding an quadratic effect of 'impluse control' on 'anxiety' with LMS, but not with the unconstrained approach. The additional quadratic effect might be spurious. In the Appendix, we show how the original and extended unconstrained approach can be implemented in the sem package in R. This gives the opportunity for applied researches to estimate latent nonlinear effects within a noncommercial software.

There are several advantages of estimating multiple nonlinear effects. One advantage refers to the development of behavioral theories containing interaction and quadratic effects (Ganzach, 1997). For example, Ganzach's theory on educational expectations leads to completely different predictions when models are estimated that contain both effect types, interaction and quadratic effects, instead of containing one interaction effect only. A simple interaction effect model predicts a high educational expectation, if both parents' educational levels are high, whereas Ganzach's theory hypothesizes a compensatory effect of the parents' educational levels. Therefore, models with an adequate amount of nonlinear effects need to be estimated and need to be accessible for a broad audience.

Another advantage results from a statistical perspective. A model with both effect types (i.e., with one interaction and two quadratic efects) serves better as a comparison model than a linear one when testing the significance of interaction effects with the χ^2-difference test (Klein, Schermelleh-Engel, Moosbrugger, & Kelava, 2009). In contrast to the hitherto widespread usage of the linear model as a comparison model for interaction effects, Klein et al. argue that an additive model containing quadratic and linear effects is more adequate. One important point is that spurious interaction effects can occur instead of true - but unspecified - quadratic effects due to the correlation of the nonlinear terms, if the linear predictors are correlated (cf. Ganzach, 1997; Lubinsky & Humphreys, 1990).

Therefore the extension of the unconstrained approach for the simultaneous estimation of quadratic and interaction effect is an important issue for testing the significance of interaction effects.

There are some limitations that need to be considered. First, we concentrated on the usage of non-overlapping indicators for each nonlinear term (e.g. x_1x_4, x_2x_5, and x_3x_6 as indicators of $\xi_1\xi_2$, instead of using $x_1x_4, x_1x_5, x_1x_6, x_2x_5$ etc.). This was necessary in order to reduce the model complexity and might lead to a slight decrease in validity. But, a modification with different numbers of indicators can be implemented with a reasonable amount of effort (for a comparison of different numbers of nonlinear indicators for the interaction model see Marsh et al., 2004). Second, although the approach has proven to be robust under some circumstances, an underestimation of standard errors of the nonlinear effects can occur which leads to an increased Type I error rate if assumptions are violated (e.g. nonnormal distribution of the linear predictors, high multicollinearity). It might be an advantage to use bootstrap procedures to overcome this problem (Efron, 1979).

REFERENCES

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

Ajzen, I. (1987). Attitudes, traits, and actions: Dispositional prediction of behavior in personality and social psychology. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 1-63). New York: Academic Press.

Algina, J., & Moulder, B. C. (2001). A note on estimating the Jöreskog-Yang model for latent variable interaction using LISREL 8.3. *Structural Equation Modeling, 8*, 40-52.

Arminger, G., & Muthén, B. O. (1998). A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis-Hastings algorithm. *Psychometrika, 63*, 271-300.

Aroian, L. A. (1944). The probability function of the product of two normally distributed variables. *The Annals of Mathematical Statistics, 18*, 265-271.

Bohrnstedt, G. W., & Marwell, G. (1978). The reliability of the products of two random variables. In K. Schüssler (Ed.), *Sociological methodology* (pp. 254-273). San Francisco: Jossey-Bass.

Bollen, K. A. (1995). Structural equation models that are nonlinear in latent variables: A least squares estimator. *Sociological Methodology, 1995*, 223-251.

Brandt, H. (2009). *Der verteilungskorrigierte Bootstrap zur Schätzung von Konfidenzintervallen nichtlinearer Strukturgleichungsmodelle* [The distribution-corrected bootstrap for the estimation of confidence intervals in nonlinear structural equation models]. Unpublished master's thesis, Institute of Psychology, Goethe University Frankfurt.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions.* New York: Irvington.

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, Ser. B, 39*, 1-38.

Diestel, S., & Schmidt, K.-H. (2009). Mediator and moderator effects of demands on self-control in the relationship between work load and indicators of job strain. *Work & Stress, 23,* 60-79.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics, 7,* 1-26.

Fox, J. (2006). Structural equation modeling with the sem package in R. *Structural Equation Modeling, 13,* 465-486.

Ganzach, Y. (1997). Misleading interaction and curvilinear terms. *Psychological Methods, 3,* 235-247.

Jaccard, J., & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin, 117*, 348-357.

Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's Reference Guide* [Computer software manual]. Lincolnwood, IL: Scientific Software International.

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

Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implication for job redesign. *Administrative Quarterly, 24,* 285-307.

Kelava, A. (2009). *Multikollinearität in nicht-linearen latenten strukturgleichungsmodellen* [Multicollinearity in non-linear structural equation models]. Unpublished doctoral dissertation, Goethe University Frankfurt.

Kelava, A., Moosbrugger, H., Dimitruk, P., & Schermelleh-Engel, K. (2008). Multicollinearity and missing constraints: A comparison of three approaches for the analysis of latent nonlinear effects. *Methodology, 4,* 51-66.

Kelava, A., Werner, C., Schermelleh-Engel, K., Moosbrugger, H., Zapf, D., Ma, Y., et al. (under revision). Advanced nonlinear structural equation modeling: Theoretical properties and empirical application of the distribution-analytic LMS and QML estimators. *Structural Equation Modeling.*

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

Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika, 65* , 457-474.

Klein, A. G., & Muth_en, B. O. (2007). Quasi maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research, 42,* 647-674.

Klein, A. G., Schermelleh-Engel, K., Moosbrugger, H., & Kelava, A. (2009). Assessing spurious interaction effects. In T. Teo & M. Khine (Eds.), *Structural equation modeling in educational research: Concepts and applications* (pp. 13-28). Rotterdam, NL: Sense Publishers.

Lee, S. Y., Song, X. Y., & Poon, W. Y. (2004). Comparison of approaches in estimating interaction and quadratic effects of latent variables. *Multivariate Behavioral Research, 39*, 37-67.

Lee, S. Y., Song, X. Y., & Tang, N. S. (2007). Bayesian methods for analyzing structural equation models with covariates, interaction, and quadratic latent variables. *Structural Equation Modeling, 14*, 404-434.

Little, T. D., Bovaird, J. A., & Widaman, K. F. (2006). On the merits of orthogonalizing powered and interaction terms: Implications for modeling interactions among latent variables. *Structural Equation Modeling, 13*, 497-519.

Lubinsky, D., & Humphreys, L. G. (1990). Assessing spurious moderator effects: Illustrated substantively with the hypothesized (synergistic) relation between spatial and mathematical abilities. *Psychological Bulletin, 107*, 385-393.

Lusch, R. F., & Brown, J. R. (1996). Interdependency, contracting, and relational behavior in marketing channels. *Journal of Marketing, 60*, 19-38.

MacCallum, R. C., & Mar, C. M. (1995). Distinguishing between moderator and quadratic effects in multiple regression. *Psychological Bulletin, 118,* 405-421.

Marsh, H. W., Wen, Z., & Hau, K.-T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods, 9*, 275-300.

Marsh, H. W., Wen, Z., & Hau, K.-T. (2006). Structural equation models of latent interaction and quadratic effects. In G. R. Hancock & R. O. Müller (Eds.), *Structural equation modeling: A second course.* (pp. 225{265). Greenwich,CT: Information Age Publishing.

Moosbrugger, H., Schermelleh-Engel, K., Kelava, A., & Klein, A. G. (2009). Assessing spurious interaction effects. In T. Teo & M. Khine (Eds.), *Structural equation modeling in educational research: Concepts and applications* (pp. 103-135). Rotterdam, NL: Sense Publishers.

Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide.* Los Angeles, CA: Muthén & Muthén.

Ping, R. A. (1995). A parsimonious estimating technique for interaction and quadratic latent variables. *Journal of Marketing Research, 32*, 336-347.

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

R Development Core Team. (2008). *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria. Available from http://www.R-project.org (ISBN 3-900051-07-0)

Schermelleh-Engel, K., Klein, A. G., & 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.

Schumacker, R. E., & Marcoulides, G. A. (1998). *Interaction and nonlinear effects in structural equation modeling.* Mahwah, NJ: Lawrence Erlbaum Associates.

Snyder, M., & Tanke, E. D. (1976). Behavior and attitude: Some people are more consistent than others. *Journal of Personality, 44*, 501-517.

Wall, M. M., & Amemiya, Y. (2000). Estimation for polynomial structural equation models. *Journal of the Statistical American Association, 95,* 929-940.

Wall, M. M., & Amemiya, Y. (2001). Generalized appended product indicator procedure for nonlinear structural equation analysis. *Journal of Educational and Behavioral Statistics, 26*, 1-29.