Quality & Quantity (2008) 42:369–396
DOI 10.1007/s11135-006-9050-6
© Springer 2006

Simple, Efficient and Distribution-free Approach to Interaction Effects in Complex Structural Equation Models

GERMÀ COENDERS[1,*], JOAN M. BATISTA-FOGUET[2] and WILLEM E. SARIS[2,3]
[1]*Department of Economics, Faculty of Economics and Management, University of Girona, Montilivi Campus, E-17071, Girona, Spain*
[2]*ESADE, Ramon Llull University, Barcelona, Spain*
[3]*Department of Sociology, University of Amsterdam, Amsterdam, The Netherlands*

Abstract. Structural equation models with mean structure and non-linear constraints are the most frequent choice for estimating interaction effects when measurement errors are present. This article proposes eliminating the mean structure and all the constraints but one, which leads to a more easily handled model that is more robust to non-normality and more general as it can accommodate endogenous interactions and thus indirect effects. Our approach is compared to other approaches found in the literature with a Monte Carlo simulation and is found to be equally efficient under normality and less biased under non-normality. An empirical illustration is included.

Key words: Interaction effects; structural equation models; indirect effects; non-linear constraints; mean structure; normality assumption.

1. Introduction

Moderated regression analysis (MRA)–a particular specification of multiple linear regression analysis which includes products of regressors–has been widely used in the social sciences when the value of a continuous variable influences the effect of another continuous variable on the dependent one (e.g. Cohen and Cohen, 1983; Jaccard et al., 1990; Irwin and McClelland, 2001; Conner et al., 2002; Davis-Blake et al., 2003; Newsom et al., 2003). However, measurement error makes the estimates of regression coefficients in MRA inconsistent and biased. This bias is especially relevant for interaction effects that are usually of low magnitude (second

*Author for correspondence: Germà Coenders, Department of Economics, Faculty of Economics and Management, University of Girona, Montilivi Campus, E-17071, Girona, Spain. E-mail: germa.coenders@udg.es

order effects) and may easily go undetected if attenuated. Additionally, the estimated standard errors of regression coefficients are also biased; thus no coherent inferences about population parameters or relationships among variables can be made.

Kenny and Judd (1984) proposed a possible specification for modeling interaction effects with structural equation models (SEM). Kenny and Judd's approach requires each latent variable to relate to at least two indicators and implies the formation of multiple indicators based on the products of the observed variables. These products are then used as indicators of the latent interaction.

Different alternatives have been proposed for developing Kenny and Judd's (1984) approach. It is not our aim to provide a comprehensive presentation of the various available procedures for testing interaction effects with SEM (see for this purpose Jöreskog, 1998; Li et al., 1998; Schumacker and Marcoulides, 1998; Cortina et al., 2002), we only cluster them into:

1. Two-step approaches (Mathieu et al., 1992; Ping, 1995, 1996; Schumacker, 2002; Bisbe et al., 2006; see also Jöreskog, *Latent Variable Scores and Their Uses*, at http://www.ssicentral.com/lisrel/techdocs/lvscores.pdf). In a first run, based on the main effects' indicators, they estimate certain parameter values of the measurement model or alternatively factor scores that are later used for the interaction factor in the second run. They have the disadvantage that quantities that are estimated in the first run are treated as known in the second, which biases standard errors. On the positive side, some of these approaches are the only feasible ones when the sample size is small (Bisbe et al., 2006) or the model is very large (Gonzalez et al., in press).
2. Two-stage least squares approach (Bollen and Paxton, 1998). It has the disadvantage of using a limited information estimator, which leads to less efficient estimates (Schermelleh-Engel et al., 1998; Moulder and Algina, 2002).
3. Maximum likelihood approaches of Jaccard and Wan (1995, 1996) and of Jöreskog and Yang (1996). Both approaches estimate the complete model with the main effects, interaction term and measurement part in one step. Both require the specification of non-linear constraints, some of which are heavily based on the normality of main effect indicators. Jöreskog and Yang use only one product indicator of the latent interaction and a mean and covariance structure, while Jaccard and Wan use all possible products of pairs of indicators and a covariance structure. Due to the use of only one indicator, the Jöreskog and Yang approach shares typical features of limited information approaches if the indicators are not congeneric or the sample size is small, and the

results can change depending on which indicator is chosen (Saris et al., in press). As a compromise, Batista-Foguet et al. (2004a) and Marsh et al. (2004) independently suggested using non-overlapping pairs of indicators of the latent interaction, thus moving to a truly full information approach in comparison with Jöreskog and Yang (1996) and leading to a simpler specification without correlated errors in comparison with Jaccard and Wan (1995). This non-overlapping pairs of indicators approach does not require non-linear constraints for identification and is thus simpler and can be applied to non-normal main effect indicators.

4. Maximum likelihood approach without product indicators or non-linear constraints (Klein and Moosbrugger, 2000; Schermelleh-Engel et al., 1998) This approach was named latent moderated structural equations (LMS) approach and does not imply creating or constraining any product indicators. It derives the joint distribution of the main effect indicators and dependent variable indicators assuming normality of the former, and expresses the likelihood as a function of all model parameters, including interaction effects. Estimation is computationally very intensive and is carried out by means of the EM algorithm as described in Klein and Moosbrugger (2000). This approach is reported to be unbiased and efficient under normality (Schermelleh-Engel et al., 1998) but biased under non-normality (Klein and Moosbrugger, 2000).
5. Some recent alternative approaches include the Bayesian approach of Arminger and Muthén (1998) the two-step method of moments approach of Wall and Amemiya (2000), and the quasi maximum likelihood estimation method (Klein and Muthén, Quasi maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Submitted to Multivariate Behavioral Research*). none of which has been implemented in widely used software thus far.

Since the newest developments not implemented in software are somewhat remote from practitioners' interests and the two-step and two-stage least squares approaches make it difficult to compute either most of the diagnosis indexes commonly used in SEM or correct standard errors, this article focuses on the maximum likelihood strategies 3 and 4.

Although the seminal article of Kenny and Judd (1984) was published 20 years ago, SEM applications to modeling interactions in the social sciences are rare (for an exception see Reinecke 2002) and MRA is mostly used instead, despite the latter's obvious weakness in not taking measurement errors into account. This indicates that cumbersome approaches (complex non linear constraints) have deterred practitioners from using SEM. This is hardly surprising, given that these approaches

require a great deal of statistical expertise. Furthermore, not all SEM software can handle non-linear constraints, and software handling the LMS approach has appeared only very recently, as well as Batista-Foguet et al. (2004a) and Marsh et al. (2004) articles. Moreover, probably due to the influence of the MRA approach, and to the complexity of the non-linear constraints, every SEM application (except that of Batista-Foguet et al. 2004b) is restricted to a single equation model, thus limiting SEM to only estimating direct effects.

This article has two main objectives: first, to extend the single equation approach to a simultaneous structural equation system in which the main effect and interaction terms may not be exogenous and in which variables may be related through a multitude of direct and indirect effects; Second to make the approach more workable and dependent on fewer assumptions by using only a few simple constraints. Going into detail, we achieve both objectives by:

1. Following Batista-Foguet et al. (2004a) and Marsh et al. (2004) in simplifying Jaccard and Wan (1995) and Jöreskog and Yang's (1996) approaches by avoiding most of the non-linear constraints and by using non-overlapping pairs of indicators of the latent interaction. This makes it possible to extend the applicability of the method to variables that are not normally distributed.
2. Following Jaccard and Wan (1995) approach by avoiding the need to use a mean structure. This results in a further simplification of the Marsh et al. (2004) approach.
3. Generalizing the single equation structural model to a simultaneous structural equation system. This would allow the researcher to have the interaction anywhere in the model, making it possible to estimate direct, indirect and total effects of the latent variables involved, thus exploiting the full strength of SEM. The earlier attempt of Batista-Foguet et al. (2004b) would be impracticable for models of moderate complexity.

The plan of this article is as follows. First we introduce the classic approaches of Jaccard and Wan (1995) and Jöreskog and Yang (1996). Then we proceed to respecify the model with the minimum possible amount of constraints, by first dropping those that are only correct under normality. Next we show the extension of the approach to models where interacting variables are endogenous and may be related to the dependent variables through indirect effects. Next, the practical feasibility of our approach is assessed in terms of bias and efficiency by means of a Monte-Carlo experiment in which our approach is compared to more complex SEM specifications and to the LMS and Marsh et al. approaches. Finally, an application is presented studying indirect and

interaction effects of political orientation, political understanding, value orientations with respect to the environment and political influence on environmentally friendly behavior such as buying and boycotting certain products.

2. Classic Single-Equation Formulations of Jöreskog and Yang and Jaccard and Wan

Before we present our proposals for extending and simplifying the approach, we start with the standard single equation SEM (Figure 1) that is presented in Equation 1:[1]

$$\eta_4 = \alpha_4 + \beta_{41}\eta_1 + \beta_{42}\eta_2 + \beta_{43}\eta_3 + \zeta_4, \tag{1}$$

where $\eta_3 = \eta_1\eta_2$, β_{41},β_{42}, and β_{43}stand for the regression coefficients and ζ_4 is the disturbance term, which is assumed to be independent of (and not just uncorrelated with) η_1 and η_2.In this approach all three variables η_1, η_2, and η_3 are exogenous variables with free variances ψ_{11}, ψ_{22}, and ψ_{33}, and covariances ψ_{21}, ψ_{31}, and ψ_{32}. Even if η_1 and $\eta_2$are centered, η_3 is not. So, $E(\eta_3) = \alpha_3$ is a parameter of the model. Furthermore we have $\mathrm{Var}(\zeta_4) = \psi_{44}$.

We assume, without loss of generality, that the main effect latent variables and the endogenous latent variable have two indicators, which is enough for identification. A larger number of indicators would of course be possible.

As regards the measurement part of the model, in this article we will only consider non-overlapping pairs of interaction indicators. If we choose

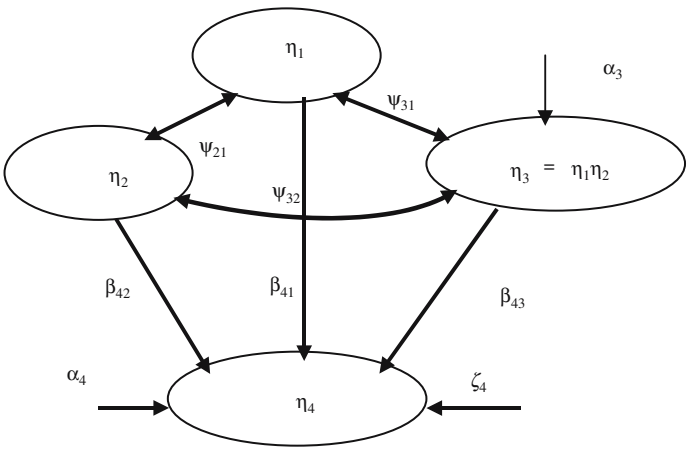

Figure 1. Single equation SEM for modeling interaction effects.

y_5=y_1y_3 and y_6=y_2y_4, then we omit the other possible indicators y_1y_4 and y_2y_3. Marsh et al. (2004) found non-overlapping pairs of indicators to work best by constructing a first pair involving the most reliable indicators of η_1 and η_2, a second pair involving the least reliable indicators of η_1 and η_2.

$$\begin{aligned}
y_1 &= \tau_1 + \lambda_{11}\eta_1 + \varepsilon_1,\\
y_2 &= \tau_2 + \lambda_{21}\eta_1 + \varepsilon_2,\\
y_3 &= \tau_3 + \lambda_{32}\eta_2 + \varepsilon_3,\\
y_4 &= \tau_4 + \lambda_{42}\eta_2 + \varepsilon_4,\\
y_5 &= \tau_5 + \lambda_{51}\eta_1 + \lambda_{52}\eta_2 + \lambda_{53}\eta_3 + \varepsilon_5,\\
y_6 &= \tau_6 + \lambda_{61}\eta_1 + \lambda_{62}\eta_2 + \lambda_{63}\eta_3 + \varepsilon_6,\\
y_7 &= \tau_7 + \lambda_{74}\eta_4 + \varepsilon_7,\\
y_8 &= \tau_8 + \lambda_{84}\eta_4 + \varepsilon_8.
\end{aligned} \tag{2}$$

For identification purposes and without any loss of generality, we fix the scale of the latent variables by constraining three loadings to 1:

$$\lambda_{11} = \lambda_{32} = \lambda_{74} = 1. \tag{3}$$

Additional parameters of the measurement part are $\mathrm{Var}(\varepsilon_j) = \theta_{jj}$. The specification is completed with the assumptions that η_1,η_2 ,and $\varepsilon_1$to ε_8 have zero expectation. Additionally, $\varepsilon_1$to $\varepsilon_8$are assumed to be mutually independent (not only uncorrelated) and independent of η_1,η_2, and ζ_4.

These assumptions allow us to analyze the expectation, variance and covariance of the product indicators, as well as to derive non-linear constraints relating their associated parameters.

The product indicators,y_5, y_6, can be decomposed as:

$$\begin{aligned}
y_5 &= y_1y_3 = (\tau_1 + \eta_1 + \varepsilon_1)(\tau_3 + \eta_2 + \varepsilon_3)\\
&= \tau_1\tau_3 + \tau_3\eta_1 + \tau_1\eta_2 + \eta_2\eta_1 + \varepsilon_5,
\end{aligned} \tag{4}$$

$$\begin{aligned}
y_6 &= y_2y_4 = (\tau_2 + \lambda_{21}\eta_1 + \varepsilon_2)(\tau_4 + \lambda_{42}\eta_2 + \varepsilon_4)\\
&= \tau_2\tau_4 + \tau_4\lambda_{21}\eta_1 + \tau_2\lambda_{42}\eta_2 + \lambda_{21}\lambda_{42}\eta_2\eta_1 + \varepsilon_6,
\end{aligned} \tag{5}$$

where:

$$\begin{aligned}
\varepsilon_5 &= \tau_3\varepsilon_1 + \tau_1\varepsilon_3 + \eta_1\varepsilon_3 + \eta_2\varepsilon_1 + \varepsilon_1\varepsilon_3,\\
\varepsilon_6 &= \tau_4\varepsilon_2 + \tau_2\varepsilon_4 + \lambda_{21}\eta_1\varepsilon_4 + \lambda_{42}\eta_2\varepsilon_2 + \varepsilon_2\varepsilon_4.
\end{aligned} \tag{6}$$

The following constraints on the loadings and intercepts can be derived from the expressions of $y_5$and$y_6$in Equations 2,4 and 5:

$$\begin{aligned}
\tau_5 &= \tau_1\tau_3,\\
\lambda_{51} &= \tau_3,\\
\lambda_{52} &= \tau_1,\\
\lambda_{53} &= 1,
\end{aligned} \tag{7}$$

$$\begin{aligned}
\tau_6 &= \tau_2\tau_4,\\
\lambda_{61} &= \tau_4\lambda_{21},\\
\lambda_{62} &= \tau_2\lambda_{42},\\
\lambda_{63} &= \lambda_{21}\lambda_{42}.
\end{aligned} \tag{8}$$

The measurement error variances and covariances can also be derived, which involves the following constraints:

$$\begin{aligned}
\theta_{51} &= E(\varepsilon_5\varepsilon_1) = \tau_3 E(\varepsilon_1\varepsilon_1) = \tau_3\theta_{11},\\
\theta_{53} &= E(\varepsilon_5\varepsilon_3) = \tau_1 E(\varepsilon_3\varepsilon_3) = \tau_1\theta_{33},
\end{aligned} \tag{9}$$

$$\begin{aligned}
\theta_{62} &= E(\varepsilon_6\varepsilon_2) = \tau_4 E(\varepsilon_2\varepsilon_2) = \tau_4\theta_{22},\\
\theta_{64} &= E(\varepsilon_6\varepsilon_4) = \tau_2 E(\varepsilon_4\varepsilon_4) = \tau_2\theta_{44},
\end{aligned} \tag{10}$$

while the error terms of non-overlapping pairs of product indicators are uncorrelated:

$$\begin{aligned}
\theta_{65} = E[&(\tau_4\varepsilon_2 + \tau_2\varepsilon_4 + \lambda_{21}\eta_1\varepsilon_4 + \lambda_{42}\eta_2\varepsilon_2 + \varepsilon_2\varepsilon_4)\\
&(\tau_3\varepsilon_1 + \tau_1\varepsilon_3 + \eta_1\varepsilon_3 + \eta_2\varepsilon_1 + \varepsilon_1\varepsilon_3)] = 0.
\end{aligned} \tag{11}$$

The error variances can also be expressed as functions of other parameters, as:

$$\begin{aligned}
\theta_{55} = Var(\varepsilon_5) &= Var(\tau_3\varepsilon_1 + \tau_1\varepsilon_3 + \eta_1\varepsilon_3 + \eta_2\varepsilon_1 + \varepsilon_1\varepsilon_3)\\
&= \tau_3^2 Var(\varepsilon_1) + \tau_1^2 Var(\varepsilon_3) Var(\eta_1) Var(\varepsilon_3)\\
&\quad + Var(\eta_2) Var(\varepsilon_1) + Var(\varepsilon_1) Var(\varepsilon_3)\\
&= \tau_3^2\theta_{11} + \tau_1^2\theta_{33} + \psi_{11}\theta_{33} + \psi_{22}\theta_{11} + \theta_{11}\theta_{33}
\end{aligned} \tag{12}$$

$$\begin{aligned}
\theta_{66} = Var(\varepsilon_6) &= Var(\tau_4\varepsilon_2 + \tau_2\varepsilon_4 + \lambda_{21}\eta_1\varepsilon_4 + \lambda_{42}\eta_2\varepsilon_2 + \varepsilon_2\varepsilon_4)\\
&= \tau_4^2\theta_{22} + \tau_2^2\theta_{44} + \lambda_{21}^2\psi_{11}\theta_{44} + \lambda_{42}^2\psi_{22}\theta_{22} + \theta_{22}\theta_{44}.
\end{aligned} \tag{13}$$

The expectation of the interaction latent variable will be:

$$\alpha_3 = E(\eta_3) = E(\eta_1\eta_2) = Cov(\eta_1\eta_2) = \psi_{21} \tag{14}$$

Some constraints on the interaction factor variances and covariances are possible if the normality assumption is made besides the independence assumption (Anderson, 1984; Jaccard and Wan, 1995; Rigdon et al., 1998):

$$\begin{aligned}\psi_{33} &= Var(\eta_3) = Var(\eta_1\eta_2) = Var(\eta_1)Var(\eta_2) + Cov^2(\eta_1\eta_2)\\ &= \psi_{11}\psi_{22} + \psi_{21}^2,\end{aligned} \tag{15}$$

$$Cov(\eta_3\eta_2) = \psi_{32} = 0;\ Cov(\eta_3\eta_1) = \psi_{31} = 0. \tag{16}$$

Jaccard and Wan (1995) suggested ignoring the mean structure. This involves centering the y_1 to y_4 indicators prior to computing the product indicators, and using only the covariance matrix as input for the model estimation (i.e., recentering the interaction indicators once they have been computed). Of course, if the α and τ parameters are of substantive interest to the researcher, then the mean structure is necessary, but this is often not the case in applied research. Some advantages of using centered indicators are:

1. Frequent non-convergence problems occur when the main effect indicators are not centered (Algina and Moulder, 2001).
2. Centering variables prior to computing the product minimizes the relationships between the variables and the product computed from them, which reduces collinearity (See Li et al. 1998 and the appendix in Irwin and McClelland 2001).
3. The number and complexity of non-linear constraints gets greatly reduced.

If we omit the mean structure, then all τ and α parameters become zero and many restrictions are no longer needed and others are greatly simplified. Constraints 7–10 on the loadings and error covariances thus simplify to $\lambda_{51} = \lambda_{52} = \lambda_{61} = \lambda_{62}$=0 and $\theta_{51} = \theta_{53} = \theta_{62} = \theta_{64}$=0. Only the following constraints on the loadings, error variances and interaction variances and covariances remain:

$$\lambda_{63} = \lambda_{21}\lambda_{42}, \tag{17}$$

$$\theta_{55} = \psi_{11}\theta_{33} + \psi_{22}\theta_{11} + \theta_{11}\theta_{33}, \tag{18}$$

$$\theta_{66} = \lambda_{21}^2\psi_{11}\theta_{44} + \lambda_{42}^2\psi_{22}\theta_{22} + \theta_{22}\theta_{44}, \tag{19}$$

$$\psi_{33} = \psi_{11}\psi_{22} + \psi_{21}^2, \tag{20}$$

$$\psi_{32} = \psi_{31} = 0. \tag{21}$$

Jaccard and Wan (1995) suggested dropping the mean structure and using all constraints (17)–(21). Additionally, they introduced all other possible pairs of product indicators (y_1y_4 and y_2y_3) with their corresponding additional constraints analogous to (17)–(19). The error covariances between overlapping pairs ($y_1y_4 - y_1y_3$; $y_1y_4 - y_2y_4$; $y_2y_4 - y_2y_3$; $y_1y_3 - y_2y_3$) were left free.

Jöreskog and Yang (1996) took a different path to simplifying the constraints. They argued that the mean structure was necessary because of the α_3 and α_4 parameters and they kept it while using only one indicator (y_5). As all constraints related to parameters involving y_6 are not needed, only constraints (7), (9), (12), (14)–(16) remain. Some of these constraints constitute what makes it possible for a model with only one indicator of one factor to be exceptionally identified.

Algina and Moulder (2001) employed much the same approach as Jaccard and Wan (1995) with two differences: They derived appropriate non-linear constraints for the error covariances between overlapping pairs, and they centered the main effect indicators prior to computing the interaction indicators but did not recenter the interaction indicators after their computation. It must be noted that the interaction indicators do not have a zero mean even if the main effect indicators do. Accordingly, they included a mean structure in which τ_1 to τ_4 were constrained to zero. The problem with these constraints is that, for centered data, they hold not only for the population but also for the sample and can thus lead to overfitting the model to the sample. Constraint (14) was also included.

Regardless of whether we take the approach of Jöreskog and Yang (1996), of Jaccard and Wan (1995) or of Algina and Moulder (2001), estimating the model with all these non-linear constraints still requires a considerable technical capability of the researcher and it is even not possible with software that does not include non-linear constraints. This is certainly one of the reasons why this approach is not frequently used in applied research. Besides, some of these constraints assume normality.

Fortunately, the estimation can be done with fewer or even no constraints at all. Marsh et al. (2004) modified the approach of Algina and Moulder (2001) by dropping overlapping pairs of indicators and by using no constraints at all except (14) which is in fact linear. The treatment of the mean structure was done as in Algina and Moulder (2001) and included setting τ_1 to $\tau_4$to 0 and thus can supposedly lead to overfitting problems. This approach is further improved on in the next section.

3. Specification with Minimal Constraints that are Robust to Non-Normality

Constraints (20) and (21) on the interaction variances and covariances make use of a result regarding the variances and covariances of products of normal variables. If the original measures are severely non-normal, then the actual variances and covariances can be very different from the values implied by the constraints (Rigdon et al., 1998) and estimates can be biased (Wall and Amemiya, 2001).

Fortunately, constraints (20) and (21) are not necessary for identification, even in the single indicator case. Researchers can deal with non-normality merely by relaxing these constraints and thus, the cost of non-normality is no more than a loss of parsimony (Rigdon et al., 1998). However, practitioners customarily introduce constraints (20) and (21) without performing any normality test.

In this article, we follow Marsh et al. (2004) and Batista-Foguet et al. (2004a) in suggesting using multiple non-overlapping pairs of indicators. Using overlapping pairs would only add unnecessary complexity in exchange of information that is actually redundant, as shown by the many error covariance parameters in Jaccard and Wan's formulation. Using only one indicator would still require the complex constraint (18) on the error variance for identification.

In the non-overlapping indicator case, no constraints are needed for identification. The issue here is one of bias and efficiency. Introducing wrong constraints on the interaction variances and covariances ((20) and (21) for non-normal data) can lead to bias. Introducing correct constraints on the error variances and loadings (17–19) can reduce standard errors. Thus, we advise practitioners to use the very simple non-linear constraint (17) on the loadings if software permits. Constraints (18) and (19) do not require normality and may also be used by expert modelers. Constraints (20) and (21) require normality and are thus not advised.

We depart from Marsh et al. (2004) in that:

1. We still consider the simple constraint (17) on the loadings while Marsh et al. suggested using no non-linear constraints at all.
2. We introduce the further simplification of omitting the mean structure completely as Jaccard and Wan (1995), while Marsh et al. suggested keeping it with Equation 14 and the main effect τ parameters constrained to zero as in Algina and Moulder (2001). This makes our approach even simpler and prevents us from overfitting the model to the sample.
3. We consider simultaneous equations and indirect effects, as shown next.

4. Extension to a Simultaneous Structural Equation System

With the exception of Batista-Foguet et al. (2004b), SEM have so far been formulated with one equation, where the regressors that interact are exogenous latent variables. This single equation formulation only makes it possible to estimate direct and interaction effects. In order to also estimate for instance an indirect effect in the same model, we must specify a simultaneous structural equation system. We show a very simple example in Figure 2.

The structural part of the model would include Equation 1 (without the α_4 term) and an additional equation predicting the up to now exogenous variable η_2.

$$\eta_2 = \beta_{21}\eta_1 + \zeta_2. \tag{22}$$

Only η_1 and $\eta_3 = \eta_1\eta_2$ are treated as exogenous. Additional parameters of the structural part are: ψ_{11}, ψ_{22}, ψ_{33}, ψ_{44}, ψ_{32} and ψ_{31}.We must take into account that $\psi_{32}$is not the covariance between η_2 and η_3 but between ζ_2 and η_3 . The assumptions are the same as before with the addition that ζ_2 is uncorrelated with η_1 (i.e. $\psi_{21} = 0$).

The model in this section also includes the relationship between η_1 and η_2 and thus both a direct and an indirect effect from η_1 to η_4. The main difference with the previous specification is that not all the variances and covariances of η_1 and η_2 are model parameters but rather functions of model parameters that can be derived from path analysis or from variance and covariance algebra (Batista-Foguet et al., 2004b):

$$Var(\eta_2) = \psi_{22} + \beta_{21}^2\psi_{11}, \tag{23}$$

$$Cov(\eta_1\eta_2) = \beta_{21}\psi_{11}. \tag{24}$$

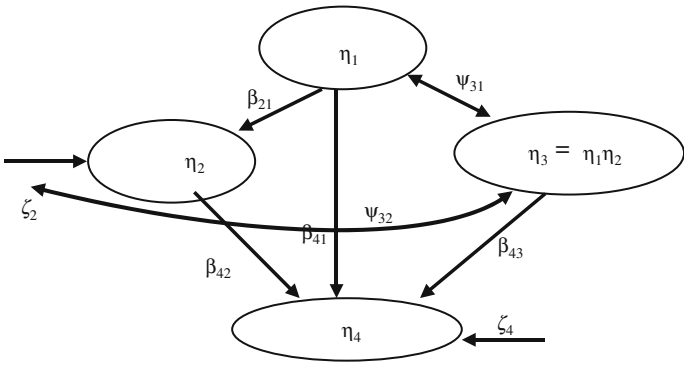

Figure 2. SEM including direct, indirect and interaction effects simultaneously.

Substituting in Equations 17–21 the new expression for Var(η_2) and Cov(η_1,η_2) yields the new constraints that apply to this model:

$$\lambda_{63} = \lambda_{21}\lambda_{42}, \tag{25}$$

$$\theta_{55} = \psi_{11}\theta_{33} + (\beta_{21}^2\psi_{11} + \psi_{22})\theta_{11} + \theta_{11}\theta_{33}. \tag{26}$$

$$\theta_{66} = \lambda_{21}^2\psi_{11}\theta_{44} + (\beta_{21}^2\psi_{11} + \psi_{22})\lambda_{42}^2\theta_{22} + \theta_{22}\theta_{44} \tag{27}$$

$$\psi_{33} = \psi_{11}(\psi_{22} + \beta_{21}^2\psi_{11}) + (\beta_{21}\psi_{11})^2 \tag{28}$$

$$Cov(\eta_3\eta_2) = \psi_{32} + \beta_{21}\psi_{31} = 0, Cov(\eta_3\eta_1) = \psi_{31} = 0. \tag{29}$$

For the single indicator approach (the only one considered in Batista-Foguet et al. 2004b), constraint (26) on the error variance is the only one needed. It does not require normality, but has increased its complexity due to the fact that Var(η_2) is a function of model parameters. This complexity would grow to unbearable levels even for moderately sophisticated models. Just think of a model including exogenous variables affecting $\eta_1$and η_2, such as that used in Section 8 (see Figure 3).

For the non-overlapping-indicator approach, none of the constraints is needed. Constraint (25) on the loadings has not increased its complexity and is still recommended, software permitting. Constraints (26) and (27) on the error variances have increased their complexity as has just been said. Constraints (28) and (29) on the interaction variances and covariances have also increased their complexity and besides require normality and are thus neither required nor advised for any of the approaches.

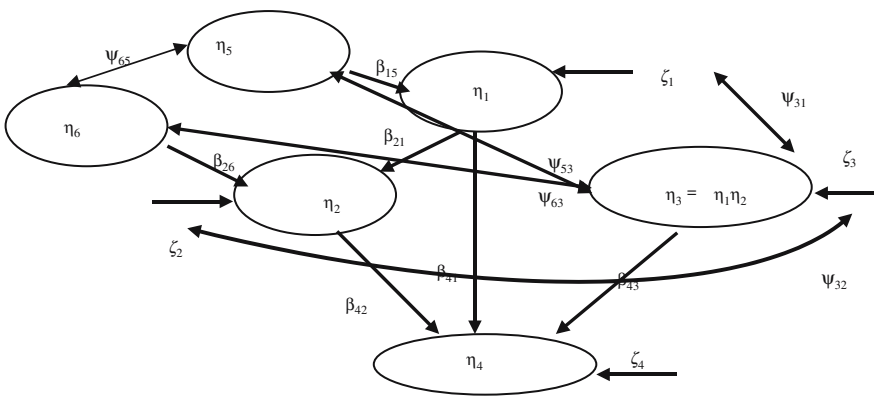

Figure 3. Model of the illustration with extra exogenous variables.

Although we illustrated this issue for a simple case, it applies to SEM with one or more interaction terms anywhere in the model, as constraint (25), the only one that we recommend for the non-overlapping indicator approach, will remain always the same. This means that by using non overlapping indicators these interaction effects can be estimated with a minimal burden of extra restrictions.

5. Estimation

Normality of the interaction indicators and of the endogenous indicators is not assumed by any of the approaches and would be difficult to fulfill, as the product of two normal variables is in general not normally distributed. Maximum likelihood estimation is still appropriate but there is a requirement for robust standard errors and test statistics (Yang-Wallentin and Jöreskog, 2001). Among these are the mean scaled χ^2 statistic (Satorra and Bentler, 1988, 1994), the $T_2^*\chi^2$ statistic (Yuan and Bentler, 2000) and robust standard errors based on a sandwich procedure (Arminger and Sobel, 1990). The LISREL8.5 (Du Toit and du Toit, 2001) and M*plus*3 (Marsh et al., 2004) programs include some or all of these robust statistics and allow the researcher to introduce non-linear constraints and are thus appropriate for all analyses. If no non-linear constraints are used, virtually any SEM software is appropriate. As a new feature, M*plus*3 includes also the LMS estimation method of Klein and Moosbrugger (2000).

Note that using the aforementioned robust statistics constitutes no safeguard against the consequences of introducing constraints (15), (16), (20), (21), (28) and (29) on interaction variances and covariances for non-normal data. Robust methods apply corrections to test statistics, not to point estimates (Satorra, 1990). Wrong constraints imply model misspecifications and as such lead to biased point estimates.

6. Interpretation and Derivation of Indirect and Total Effects

The total effect of η_2 on η_4 conditional on different values of η_1, can be obtained from the expected value in the expression (Batista-Foguet et al., 2004b):

$$E(\eta_4) = \beta_{41}\eta_1 + \beta_{42}\eta_2 + \beta_{43}\eta_1\eta_2 \tag{30}$$

as the partial derivative of E(η_4) with respect to η_2.

$$\frac{\partial E(\eta_4)}{\partial \eta_2} = \beta_{42} + \beta_{43}\eta_1. \tag{31}$$

Besides a direct main effect, this equation also displays a typical interaction effect, in which the effect of η_2 on η_4 depends on the value of

η_1. Thus, the interpretation of the β_{42} main effect parameter is that occurring when the value of the other variable is zero. As η_1 and η_2 are always mean centered, β_{42}can more easily be interpreted as the effect for the mean value of the other variable.

In order to assess the effect of η_1, its relationship to η_2 has to be taken into account, so that:

$$\frac{\partial E(\eta_2)}{\partial \eta_1} = \beta_{21}. \tag{32}$$

And the partial derivative of Equation (30), now with respect to η_1 gives us the total effect of η_1 on η_4 as:

$$\frac{\partial E(\eta_4)}{\partial \eta_1} = \beta_{41} + \beta_{42}\beta_{21} + \beta_{43}(\eta_2 + \beta_{21}\eta_1). \tag{33}$$

This equation displays a direct main effect β_{41}, an indirect effect $\beta_{42}\beta_{21}$, the interaction effect $\beta_{43}\eta_2$ and a combined interaction-indirect effect $\beta_{43}\beta_{21}\eta_1$. Thus, the effect of η_1 on η_4 also depends on the value of η_1.

7. Simulation Experiment

We designed a small Monte Carlo experiment to test the performance of our suggested approach and assess its bias and efficiency in comparison with the one-indicator, LMS and Marsh et al. approaches under different distributional conditions.

For this simulation experiment we used the model in Figure 2, assuming that η_1 , η_2, and η_4 have two indicators each. We simulated 500 samples of size n=400. The parameter values are described in Appendix A and the distributional conditions were:

1. Normal distribution of all model components.
2. η_1 and ζ_2 were distributed as a χ^2 with four degrees of freedom. We selected this distribution because it leads to sizeable but not extreme departures from normality (skewness was 1.4 for η_1 and 1.1 for η_2, kurtosis was 3 for η_1 and 1.9 for η_2).
3. η_1 and ζ_2 were distributed as a χ^2 with four degrees of freedom but R^2 for η_2 was higher than for conditions 1 and 2. We manipulated this R^2 because the fulfillment of constraints on interaction variances and covariances (28) and (29) does not only depend on the distributional shape but also on the correlation between η_1 and η_2.

We analyzed the data of the 500 replications in five different ways:

1. One interaction indicator without mean structure and with complex constraints (26), (28), and (29) that assume normality (modified Jöreskog–Yang's approach without means).
2. One interaction indicator without mean structure and with minimal complex constraints on error variances (26) that do not assume normality (Modified Jöreskog–Yang's approach without either means or normality assumption).
3. Non-overlapping indicators without mean structure and with only simple constraints (25) that do not assume normality (Our approach with constrained interaction loadings).
4. Non-overlapping indicators without mean structure or constraints (Our approach with no constraints).
5. No indicators with mean structure (LMS approach).
6. Non-overlapping indicators with centered main effect indicators and mean structure with $\tau_1 = \cdots = \tau_4 = 0$ and $\alpha_3 = \beta_{21}\psi_{11}$, which is what Equation 14 becomes due to the indirect effect. (Marsh et al.'s approach).

We did all analyses by maximum likelihood using the $T_2^*\chi^2$ statistic (Yuan and Bentler, 2000) and robust standard errors based on a sandwich procedure (Arminger and Sobel, 1990), i.e., the MLR option in the M*plus*3 program (Marsh et al., 2004). We present the following summary measures of the results regarding test statistics and parameters of the structural part of the model:

1. Bias (B): difference between the mean of point estimates and the population value, as a percentage of the population value (positive for overestimation).
2. Efficiency (SD): standard deviation of the point estimates.
3. Bias of standard errors (BSE): difference between the mean of standard error estimates and the standard deviation of point estimates, as a percentage of the standard deviation of point estimates (positive for overestimation).
4. Percentage of rejections using the χ^2-test (PR): Computed at the nominal 0.05 level. It should thus approach 5%. The χ^2-test is not available for the LMS approach.

All replications converged to admissible solutions except five negative variances under distributional condition 1 (normality) for two approaches (our approaches with and without constrained loadings), and one negative variance under distributional condition 2 (non normality with low main effect correlation) for our approach with constrained loadings. The results are displayed in Table I.

Table I. Results of the Monte Carlo experiment

Approach		Jöreskog–Yang's approach without means			Jöreskog–Yang's approach without either means or normality assumption			Our approach with constrained loadings			Our approach with no constraints			LMS approach			Marsh et al.'s approach		
Condition		B%	SD	BSE%	B%	SD	BSE%	B%	SD	BSE%	B%	SD	BSE%	B%	SD	BSE%	B%	SD	BSE%
	β_{21}	−0.5	0.061	−5.1	−0.2	0.058	−4.1	−0.3	0.056	−2.0	−0.4	0.056	−1.6	−0.3	0.056	−1.6	−0.5	0.057	−1.9
	β_{41}	−1.0	0.057	−4.9	−1.0	0.058	−3.5	0.3	0.052	1.7	0.3	0.054	0.6	−1.0	0.056	−1.1	0.3	0.053	1.5
	β_{42}	0.4	0.055	−0.4	0.4	0.055	1.8	−0.8	0.055	−3.5	−0.8	0.057	−5.1	0.4	0.053	6.1	−0.8	0.057	−5.1
	β_{43}	0.7	0.057	−8.2	1.8	0.060	−6.6	0.4	0.052	−6.2	0.4	0.053	−6.5	−0.3	0.047	−5.9	0.6	0.050	−6.5
1	ψ_{11}	−1.1	0.108	1.7	−0.4	0.101	1.1	−0.3	0.101	−5.7	−0.2	0.109	−6.8	−0.4	0.097	2.5	−0.3	0.108	−6.7
normality	ψ_{22}	−0.7	0.089	−7.2	−0.2	0.087	−7.2	−0.6	0.077	−2.3	−0.5	0.081	−1.1	−0.4	0.082	−4.2	−0.6	0.081	−1.2
	ψ_{44}	−2.0	0.060	−1.8	−2.0	0.060	−1.5	−1.6	0.056	−4.6	−1.5	0.058	−2.3	−1.5	0.057	−0.7	−1.7	0.057	−1.6
	avg.	0.9	0.069	4.2	0.9	0.068	3.7	0.6	0.064	3.7	0.6	0.067	3.4	0.6	0.064	3.1	0.7	0.066	3.5
	PR		3.6%			4.6%			4.8%			4.4%			N.A.			1.8%	
	β_{21}	7.4	0.087	−6.3	−0.3	0.069	0.6	−0.3	0.058	−6.1	−0.3	0.058	−5.7	0.1	0.070	0.0	−0.3	0.058	−4.7
	β_{41}	−3.7	0.059	−8.2	−1.3	0.061	−4.8	0.2	0.063	−1.9	0.2	0.064	−2.8	−8.0	0.057	−2.6	0.1	0.064	−2.8
2	β_{42}	14.9	0.060	−8.9	−0.9	0.067	−0.4	−1.7	0.059	−6.7	−1.6	0.059	−6.4	7.1	0.059	1.2	−1.6	0.059	−6.6
	β_{43}	5.4	0.053	−3.6	3.4	0.061	−6.9	2.0	0.050	−5.2	2.1	0.050	−5.0	8.2	0.046	−0.9	2.1	0.049	−5.1
Non normality.	ψ_{11}	12.7	0.208	−15.9	−0.2	0.123	−1.5	−0.5	0.122	−0.7	−0.5	0.127	−1.9	−0.5	0.119	0.1	−0.5	0.128	−7.0
Low correlation	ψ_{22}	5.1	0.113	−11.9	−0.7	0.094	−3.5	−0.5	0.099	−2.7	−0.3	0.103	−1.5	−1.0	0.095	−4.8	−0.3	0.103	−1.5
$\eta_1 - \eta_2$	ψ_{44}	−2.8	0.074	−2.2	−1.5	0.074	−6.1	−1.4	0.070	−6.1	−1.2	0.072	−6.9	−2.0	0.071	−5.5	−1.3	0.072	−6.9
	avg.	7.4	0.093	8.1	1.2	0.078	3.4	0.9	0.074	4.2	0.9	0.076	4.3	3.8	0.074	2.2	0.9	0.076	4.9
	PR		99.8%			5.8%			5.0%			5.0%			N.A.			3.8%	

	β_{21}	6.2	0.073	−6.3	−0.1	0.066	−2.0	0.6	0.051	−1.0	0.6	0.051	−0.8	0.2	0.065	0.0	0.6	0.051	0.0
3	β_{41}	−3.5	0.074	−6.2	0.5	0.077	−4.5	0.8	0.082	−1.1	0.8	0.084	−2.4	−9.7	0.071	−1.4	0.7	0.084	−2.5
	β_{42}	14.8	0.079	−10.2	1.6	0.089	−3.6	−0.9	0.074	−4.7	−0.7	0.074	−4.3	10.8	0.075	−0.5	−0.7	0.074	−4.6
	β_{43}	6.2	0.045	−9.4	2.5	0.050	−11.5	0.2	0.044	−12.3	0.3	0.045	−12.1	8.7	0.040	−2.7	0.4	0.044	−11.2
Non normality.	ψ_{11}	23.8	0.227	−10.8	0.0	0.111	−0.8	−0.6	.123	−2.9	−0.6	0.127	−3.5	−0.3	0.109	1.1	−0.7	0.127	−11.2
High correlation	ψ_{22}	4.1	0.067	−3.3	1.8	0.061	−0.2	−0.9	0.067	−1.8	−0.8	0.069	−1.0	−2.0	0.061	1.2	−0.8	0.069	−1.2
η_1-η_2	ψ_{44}	−3.4	0.068	6.2	0.9	0.067	2.4	−0.3	0.065	1.7	−0.3	0.065	2.6	−1.2	0.064	3.4	−0.2	0.065	−2.8
	avg.	8.9	0.090	7.5	1.0	0.074	3.6	0.6	0.072	3.6	0.6	0.073	3.8	4.7	0.069	1.5	0.6	0.073	4.8
	PR		100%			6.8%			5.8%			6.8%			N.A.			2.6%	

B bias, *SD* efficiency (standard deviation), *BSE* standard error bias, *PR* percentage of rejections using the χ^2 statistic, *Avg* Absolute average over all structural parameters.

As regards bias (B), all approaches not requiring normality (the modified Jöreskog–Yang approach without either means or normality assumption, both our approaches, and Marsh et al.'s approach) are unbiased under all conditions (average bias around or below 1%). Omitting the mean structure, thus, leads to no bias even if all variables were simulated to have means different from zero (see Appendix A). The modified Jöreskog–Yang approach without means and the LMS approach lead to biased estimates under non-normality. Bias is larger when the correlation between η_1 and $\eta_2$is higher and for the modified Jöreskog–Yang approach. This bias was also reported by Marsh et al. (2004) who used the original Jöreskog–Yang approach with means.

As regards the performance of the χ^2-test statistic, it is reasonably accurate (PR close to 5%) for nearly all approaches for which the test is available for all distributional conditions. The two only exceptions are the modified Jöreskog–Yang approach without means under non normality and Marsh et al.'s approach under all conditions. In the first case the result is due to specification errors. In the second case it is due to four constraints ($\tau_1 = \cdots = \tau_4 = 0$) that hold exactly in the sample and not only in the population. In fact, under Marsh et al.'s approach, the mean of the χ^2 statistic over the 500 replications was lower than the number of degrees of freedom by about 4 under all three distributional conditions.

As regards the accuracy of standard errors (BSE), they are markedly wrong for the modified Jöreskog–Yang approach without means under non normality. All approaches except LMS tend to underestimate standard errors but to a minor extent (by less than 5% on average).

As regards efficiency (SD), from the results of Marsh et al. (2004) we expected the inclusion of additional constraints to reduce standard errors, though this did not hold under all conditions. Even under normality, the modified Jöreskog–Yang approach with extra constraints requiring normality is no better than the same approach without these extra constraints in terms of standard deviation (SD). Thus there seems to be nothing to be gained from the normality assumption even when it holds. In contrast, our approach is slightly more efficient with the constraint on the loadings than without. Marsh et al.'s approach and or approach without constrained loadings are equally efficient and seem to be no worse than both modified Jöreskog–Yang approaches and thus the introduction of additional indicators generates degrees of freedom that seem to more than offset the omission of constraints. The conclusion that the LMS approach is more efficient than the standard SEM approaches (Schermelleh-Engel et al., 1998) does not seem to hold when comparing LMS to our approach with constrained loadings. The LMS approach is, however, more efficient than all other standard SEM approaches considered.

As a conclusion, the LMS approach and our approach with constrained loadings perform best in terms of efficiency. All approaches except the LMS approach and the modified Jöreskog–Yang approach without means are robust to non-normality. All approaches except the modified Jöreskog–Yang approach without means lead to approximately correct standard errors. The modified Jöreskog-Yang approach without either means or normality assumption and both our approaches lead to correct χ^2 statistics. Thus, our approach with non-overlapping indicators and constrained loadings (constraint (25)) would generally be considered to be the best. Our approach without constraints remains an attractive second best option for researchers not having software that is able to deal with nonlinear constraints. LMS is only appropriate when normality of the main effect variables holds. The modified Jöreskog–Yang approach without either means or normality assumption, though not so bad in performance, would be unworkable for large models. Marsh et al.'s approach has the drawback of tending to accept the model too often on the basis of the χ^2 statistic, which implies a loss of statistical power. The modified Jöreskog–Yang approach without means is clearly the worst.

8. An Illustration

We illustrate our recommended approaches with and without the factor loading constraint (25) by using one example based on data from the British pilot study of the European Social Survey carried out in 2002 (see www.europeansocialsurvey.org). Our dependent variable is "environmental friendly behavior" like buying and boycotting certain products (η_4). The explanatory variables are the possibilities of influencing events, or "political influence" (η_1) and "value orientations with respect to the environment" (η_2). Besides, some exogenous variables are considered: "understanding of politics" (η_5) and "left-right political orientation" (η_6). See Appendix B for the description of the indicators of each of these variables. We expected left-right orientation to affect value orientations (β_{26}) and understanding of politics to affect political influence (β_{15}). We expected political influence to affect value orientations (β_{21}). We also expected value orientation to affect behavior only if people believed they could influence the situation. So we expected an interaction effect of these two variables which is represented by η_3 (β_{43}). The model is depicted in Figure 3, where it can be seen that it is a generalization of Figure 2 with two extra exogenous variables. This added complexity does not alter the formulation of constraint (25). Note how the interaction factor correlates with the exogenous factors and the disturbance term of variables not affected by it. Note also that the main effects β_{41} and β_{42} have to be model parameters even if only the interaction effect is theoretically relevant (Irwin and McClelland 2001).

Sample size was 409 and we used EM imputation of missing values. The main effect indicators were non-normally distributed. The maximum absolute skewness was 0.75 for y_4 and the maximum absolute kurtosis was -0.92 for y_1. As a consequence, constraints that assume normality and the LMS approach were not appropriate.

The results are displayed in Table II. We used the M*plus*3 program with the MLR option. Appropriate input files for both the M*plus*3 and LISREL8.5 programs are in Appendix C.

Both models converged to a proper solution. Even if they are rejected by the $T_2^*\chi^2$ statistic, other fit indices [root mean squared error of approximation (RMSEA), standardized root mean squared residual (SRMR), Tucker and Lewis Index (TLI) and comparative fit index (CFI)] all suggest that the fit of both models is good enough to proceed to compare their results. Both models also yielded very similar results: left–right orientation had a

Table II. Estimates and standard errors of key parameters and goodness of fit indexes with and without the simple constraint (25) on the factor loadings

Parameter/index	With (25)	Without (25)
β_{43}	-0.081	-0.060
s.e. β_{43}	0.045	0.039
β_{42}	0.104	0.100
s.e. β_{42}	0.033	0.033
β_{41}	-0.134	-0.134
s.e. β_{41}	0.046	0.047
β_{26}	-0.730	-0.730
s.e. β_{26}	0.076	0.076
β_{21}	0.006	0.006
s.e. β_{21}	0.059	0.059
β_{15}	0.603	0.604
s.e. β_{15}	0.086	0.087
R^2 (η_4) %	10.8	10.4
R^2 (η_2) %	37.0	37.0
R^2 (η_1) %	33.4	33.3
Robust T_2^* χ^2	71.49	70.69
d.f.	44	43
p-value	0.006	0.005
RMSEA	0.039	0.040
SRMR	0.032	0.033
TLI	0.958	0.956
CFI	0.972	0.972

significant effect on value orientations (β_{26}), understanding of politics had a positive effect on political influence (β_{15}), political influence had an unexpectedly negative effect on behavior (β_{41}), value orientation had a significant effect on behavior (β_{42}) which was not moderated by political influence, as the interaction effect (β_{43}) was non significant.

9. Conclusions and Recommendations

This article provides a simplified, robust and extended approach for modeling interactions using SEM.

The main idea underlying the *simplification* stems from the fact that there are still very few SEM users modeling interaction effects. Researchers keep using MRA instead because the current SEM practice simply requires too much methodological expertise.

Since Kenny and Judd (1984), available approaches have involved two-step methods with unclear theoretical and statistical properties, limited information methods, sophisticated methods out of reach of applied researchers or maximum likelihood SEM methods with very complicated nonlinear constraints, some of which require normality which, according to our simulations, is a very unwise assumption for the sake of *robustness*. Marsh et al. (2004) proposed some important simplifications and we in some respects simplify their proposal even further (omission of the mean structure) and in some other respects extend it (inclusion of indirect effects and constraint (25)).

Completely omitting the mean structure has proven appropriate in our experiment. No bias or loss of efficiency emerges a result. On the contrary, Marsh et al.'s suggestion to keep a simple constrained mean structure on the centered data has proven to negatively affect the χ^2 statistic.

The main idea underlying the *extension* of the approaches used so far is that by restricting SEM to one-equation models, up to now researchers were actually diminishing the potentiality of SEM. SEM should also be able to cope with relatively complex models, including indirect effects as well as direct. The elimination of the need for complex constraints used in Batista-Foguet et al. (2004b) makes the approach much more workable, independent of the complexity of the model and the fulfillment of normality, while allowing applied researchers to fit these models easily.

So, a very practical recommendation is to drop the mean structure, use non-overlapping indicators for the latent interaction and:

1. For a model with a single equation, constraints (17)–(19) are relatively simple, do not require normality and can be used, but they are not needed for identification.
2. For a model with several equations, only constraint (25) on the loadings remains simple and can be used though it is not required for identification.

3. For a researcher whose software cannot handle non-linear constraints, even constraint (17) or (25) can be dropped, at the expense of getting slightly larger standard errors.

Our method has shown to be unbiased even under non-normality and at least as efficient as the alternatives, even as efficient as the LMS approach which was before thought to be superior to the SEM approaches, and more efficient than Marsh et al.'s approach while providing correct χ^2 statistics.

If the main effects have more than two indicators, it is also possible to form a larger number of non-overlapping pairs as indicators of the latent interaction, which can result in higher efficiency and even smaller rates of inadmissible solutions (Boomsma and Hoogland, 2001). As before, no constraints are needed.

Appendix A: Simulated Population Parameters for the Monte Carlo Experiment

As can be seen from table III, parameter values are the same for all distributional conditions except ψ_{22} and β_{21} for the non-normal condition with highly correlated η_1 and η_2. This results in obvious changes in the R^2-value for η_2 for this condition. However, some other model characteristics (including even R^2 and reliabilities) change due to the fact that the variances and covariances of η_3 change depending on the extent to which constraints (28) and (29) are violated, which depends on the distributional condition. The table shows both the empirical variances and covariances of η_3 and those derived from the constraints and it can be seen that deviations are quite dramatic under non-normality, especially when η_1 and $\eta_2$are highly correlated. Note that parameters relating to η_3 were not pre-specified in the simulation but resulted from the computation of the product of η_1 and η_2.

Note that the model is simulated with a mean structure (parameters τ_1 to τ_4, together with Equation 14). This is made in order to check that omitting an existent mean structure in the analysis is not harmful. For the analysis, the main effect indicators were centered on the means of each specific sample and this was done prior to the computation of the product indicators.

Appendix B: The Measurement of the Different Variables of Figure 3

In the European Social Survey pilot study, the following measures were available for the different variables of interest.

The dependent variable is an "environmental friendly behavior" variable measuring behaviors such as purchasing and boycotting of products and other items for environmental and other political reasons (η_4). This measure has been asked in two different ways, the second at the end of the sup-

Table III. Simulated parameter values

Condition	Normality	Non-normality	Non normality. High correlation $\eta_1 - \eta_2$
Distribution	All normal components	$\eta_1, \quad \zeta_2 : \rightarrow \chi^2_{4d.f.}$	$\eta_1, \quad \zeta_2 : \rightarrow \chi^2_{4d.f.}$
Model parameters			
$\theta_{11}, \ldots, \theta_{44}, \theta_{77}, \theta_{88}$	0.25	0.25	0.25
$\lambda_{21}, \ldots, \lambda_{42}, \lambda_{74}, \lambda_{84}$	1.00	1.00	1.00
$\tau_1, \ldots, \tau_4$	1.00	1.00	1.00
$\tau_7, \quad \tau_8$	0.00	0.00	0.00
ψ_{11}	1.00	1.00	1.00
ψ_{22}	0.75	0.75	0.50
ψ_{44}	0.75	0.75	0.75
β_{21}	0.50	0.50	0.71
$\beta_{41}, \quad \beta_{42}, \beta_{43}$	0.30	0.30	0.30
$\alpha_2, \quad \alpha_4$	0.00	0.00	0.00
Explained variances			
Reliability $y_1, \ldots, y_4$	0.80	0.80	0.80
Reliability y_7, y_8	0.77	0.82	0.85
$R^2(\eta_2)$	0.25	0.25	0.50
$R^2(\eta_4)$	0.43	0.55	0.64
Distribution of η_3 (not simulated, computed as $\eta_1\eta_2$)[1]			
$Var(\eta_3)$	1.25/1.25	1.9/1.25	3/1.5
$Cov(\eta_1,\eta_3)$	0/0	0.7/0	1/0
$Cov(\eta_2,\eta_3)$	0/0	0.3/0	0.7/0

[1]The first figure shows the empirical distribution. The second figure shows what constraints (28) and (29) would impose.

plementary methodological questionnaire (after approximately 45 minutes of other questions):

y_7: *How many of the four things on this card have you done during the last 12 months?*

1. *Deliberately bought certain products for political, ethical or environmental reasons.*
2. *Boycotted certain products.*
3. *Donated money.*
4. *Raised funds.*

y_8: *During the last 12 months, have you done any of the following ?* ("yes"–"no" scale; the count of "yes" answers was used)

1. *Deliberately bought certain products for political, ethical or environmental reasons.*
2. *Boycotted certain products.*
3. *Donated money.*
4. *Raised funds.*

A question was asked to measure "political influence" (η_1): *How far do you agree or disagree with the following statement ?*(in a 5-point "completely disagree" to "completely agree" scale)*: I think I can take an active role in a group that is focused on political issues*

y_1: presentation of the item at the beginning of the questionnaire.

y_2: presentation of the item at the end of the questionnaire.
"Value orientations with respect to the environment" (η_2) were measured by items of Schwartz (1997) value scale: *How much like you is this person?* (in a 6-point "very much like me" to "not like me at all" scale).

y_3: *Looking after the environment is important to him/her.*

y_4: *He/she strongly believes that people should care for nature.*
For the "understanding of politics" variable (η_5) two questions were asked:

y_9: *How far do you agree or disagree with the following statement ?*(in a 5-point "completely disagree" to "completely agree" scale):*Sometimes politics and government seem so complicated that I can't really understand what is going on.*

y_{10}: *How often do politics and government seem so complicated that you can't really understand what is going on?* (in a 5-point "never" to "frequently" scale).
For the "left–right political orientation" variable (η_6) also two questions were asked:
How much like you is this person?(in a 6-point "very much like me" to "not like me at all" scale).

y_{11}: *He/she thinks it is important that every person in the world be treated equally.*

y_{12}: *He/she believes everyone should have equal opportunities in life.*

Appendix C: LISREL85 and Mplus3 Setups for the Illustration

Bold faced input corresponds to constraint (25) and can just be dropped if this constraint is not wished.

```
LISREL set-up for the illustration
DA NI=12 NO=409
LA
y1y2y3y4y1y3y2y4y7y8y9y10y11y12
CM FI=illustration.CM
AC FI=illustration.AC
MO NY=12 NE=6 LY=FU,FI PS=SY,FI TE=DI,FR BE=FU,FI
FR LY 2 1 LY 4 2 LY 6 3 LY 8 4 LY 10 5 LY 12 6
FR BE 2 1 BE 4 1 BE 4 2 BE 4 3 BE 1 5 BE 2 6
FR PS 1 1 PS 2 2 PS 3 3 PS 4 4 PS 3 1 PS 3 2 PS 5 5 PS 6 6
FR PS 6 5 PS 5 3 PS 6 3
VA 1 LY 1 1 LY 3 2 LY 5 3 LY 7 4 LY 9 5 LY 11 6
CO LY(6,3) = LY(2,1)*LY(4,2)
OU ML
```

```
title: Mplus set-up for the illustration;
data: file is illustration.dat;
variable: names are y1y2y3y4y1y3y2y4y7y8y9y10y11y12;
analysis: type is general;
estimator=mlr;
model:
eta1 by y1@1;
eta1 by y2* (p1);
eta2 by y3@1;
eta2 by y4* (p2);
eta3 by y1y3@1;
eta3 by y2y4* (p3);
eta4 by y7@1 y8*;
eta5 by y9@1 y10*;
eta6 by y11@1 y12*;
eta1 on eta5*;
eta2 on eta1* eta6*;
eta4 on eta1* eta2* eta3*;
y1* y2* y3* y4* y1y3* y2y4* y7* y8* y9* y10* y11* y12*;
eta1* eta2* eta3* eta4* eta5* eta6*;
eta5 with eta6*;
eta3 with eta1* eta2* eta5* eta6*;
model constraint:
p3 = p2*p1;
```

Note

1. The all-y (or all-η) notation is used throughout. The exogenous and endogenous status of certain variables will change from model to model and we want to maintain a homogeneous notation for ease of comparison. Besides, this notation is needed to estimate some of these models with LISREL.

References

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.

Anderson,T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd edn. New York: Wiley.

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

Arminger, G. & Sobel, M. E. (1990). Pseudo maximum likelihood estimation of mean and covariance structures with missing data. *Journal of the American Statistical Association* 85: 195–203.

Batista-Foguet, J.M., Coenders, G. & Saris, W.E. (2004a). A parsimonious approach to interaction effects in structural equation models: an application to consumer behavior. *Working Papers d'ESADE* 183: 1–28.

Batista-Foguet, J. M., Coenders, G., Saris, W. E. & Bisbe, J. (2004b). Simultaneous estimation of indirect and interaction effects using structural equation models. *Metodološki Zvezki* 1: 163-184.

Bollen, K. A. & Paxton, P. (1998). Interactions of latent variables in structural equation models. *Structural Equation Modeling* 5: 267–293.

Boomsma, A. & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In: R. Cudeck, S. du Toit & D. Sörbom (eds.), *Structural Equation Modeling: Present and Future.* Lincolnwood, IL: Scientific Software International, pp. 139–168.

Bisbe, J., Coenders, G., Saris, W. E. & Batista-Foguet, J. M. (2006). Correcting measurement error bias in interaction models with small samples. *Metodološki Zvezki* 3: 267–287

Cohen, J. & Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* Hillsdale: Lawrence Erlbaum.

Conner, M., Sparks, P., Povey, R., James, R Shepherd, R,. & Armitage, C. J. (2002). Moderator effects of attitudinal ambivalence on attitude-behaviour relationships. *European Journal of Social Psychology* 32: 705–718.

Cortina, J. M., Chen G. & Dunlap, W. P. (2002). Testing interaction effects in LISREL: examination and illustration of the available procedures.*Organizational Research Metho*ds 4: 324–60.

Davis-Blake, A., Broschak, J. P. & George, E. (2003). Happy together? How using non standard workers affects exit, voice, and loyalty among standard employees. *Academy of Management Journal* 46: 475–485.

Du Toit, M., & du Toit, S. (2001). *Interactive LISREL, User's Guide.* Chicago, IL: Scientific Software International.

González, M., Coenders, G. & Casas, F. (in press). Using non-linear models for a complexity approach to psychological well-being. *Quality and Quantity* DOI: 10.1007/s11135-006-9032-8.

Irwin, J. R. & McClelland, G. H. (2001). Misleading heuristics and moderate regression models. *Journal of Marketing Research* 38: 100–109.

Jaccard, J., Turrisi, R. & Wan, C. K., (1990). *Interaction Effects in Multiple Regression*. Newbury Park, CA: Sage.

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

Jaccard, J. & Wan, C. K. (1996). *Lisrel Approaches to Interaction Effects in Multiple Regression.* Thousand Oaks, CA: Sage.

Jöreskog, K. G. (1998). Interactions and nonlinear modeling: issues and approaches. In: R. E. Schumacker & G. A. Marcoulides (eds.), *Interactions and Nonlinear Effects in Structural Equation Models.* Mahwah, NJ: Lawrence Erlbaum, pp. 239–250.

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*. Mahwah, NJ: Lawrence Erlbaum, pp. 57–89.

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

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

Li, F., Harmer, P., Duncan, T. E., Duncan, S. C., Acock, A. & Boles, S. (1998). Approaches to testing interactions effects using structural equation modeling methodology. *Multivariate Behavioral Research* 33: 1–39.

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.

Mathieu, J. E., Tannenbaum, S. I. & Salas, E. (1992). Influences of individual and situational characteristics on measures of training effectiveness. *Academy of Management Journal* 35: 828–847.

Moulder, B. C. & Algina, J. (2002). Comparison of methods for estimating and testing latent variable interactions. *Structural Equation Modeling* 9: 1–19.

Muthén, L. K. & Muthén, B. O. (2004). *Mplus. Statistical Analysis with Latent Variables. User's Guide* 3rd edn. Los Angeles, CA: Muthén and Muthén.

Newsom, J. T., Prigerson, H. G., Schulz R., & Reynolds, C. F. (2003). Investigating moderator hypotheses in aging research: statistical, methodological, and conceptual difficulties with comparing separate regressions. *International Journal of Aging and Human Development* 57: 119–150.

Ping, R. A. (1995). A parsimonious estimating technique for interaction and quadratic latent variables. *Journal of Marketing Research* 32: 336–47.

Ping, R. A. (1996). Latent variable interaction and quadratic effect estimation: A two step technique using structural equation analysis. *Psychological Bulletin* 119: 166–75.

Reinecke, J. (2002). Nonlinear structural equation models with the theory of planned behavior: Comparison of multiple group and latent product term analyses. *Quality and Quantity* 36: 93–112.

Rigdon, E. E., Schumacker, R. E. & Wothke, W. (1998). A comparative review of interaction and nonlinear effects modeling. In: R. E. Schumacker & G. A. Marcoulides (eds.), *Interactions and Nonlinear Effects in Structural Equation Models.* Mahwah, NJ: Lawrence Erlbaum, pp. 1–16.

Saris, W. E., Batista-Foguet, J. M. & Coenders, G. (in press): Selection of indicators for the interaction term in structural equation models with interaction. *Quality and Quantity.*

Satorra, A. (1990). Robustness issues in structural equation modeling: a review of recent developments. *Quality and Quantity* 24: 367–86.

Satorra, A. & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *ASA Proceedings of the Business and Economic Section*. Alexandria, VA: American Statistical Association, pp. 308–313.

Satorra, A. & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In: A. von Eye & C. Clogg (eds.), *Latent Variables Analysis: Applications to Developmental Research.* Thousand Oaks, CA: Sage, pp. 399–419.

Schermelleh-Engel, K., Klein, A. & Moosbrugger, H. (1998). Estimating non-linear effects using a latent moderated structural equations approach. In: R. E. Schumacker & G. A. Marcoulides (eds.), *Interactions and Nonlinear Effects in Structural Equation Models.* Mahwah, NJ: Lawrence Erlbaum, pp. 203–238.

Schumacker, R. E. (2002). Latent variable interaction modeling.*Structural Equation Modeling 9*: 40–54.

Schumacker, R. E. & Marcoulides, G. A. (1998). *Interactions and Nonlinear Effects in Structural Equation Models*. Mahwah, NJ: Lawrence Erlbaum.

Schwartz, S. H. (1997). Values and culture. In: D. Munro, S. Carr & J. Schumaker (eds.), *Motivation and Culture.* New York: Routledge, pp. 69–84.

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

Yang-Wallentin, F. & Jöreskog, K. G. (2001). Robust standard errors and chi-squares for interaction models. In: G. A. Marcoulides & R. E. Schumacker (eds.), *New Developments and Techniques in Structural Equation Modeling*, Mahwah, NJ: Lawrence Erlbaum, pp. 159–171.

Yuan, K. H. & Bentler P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. In: M. E. Sobel & M. P. Becker (eds.), *Sociological Methodology 2000*. Washington, DC: American Sociological Association, pp. 165–200.