Standardized Estimation of Latent Variable Interaction Effects: Method Comparison and Selection Strategies

Authors: Zhonglin Wen, Ouyang Jinying, Fang Junyan, Zhonglin Wen

Date: 2021-10-12T00:00:00+00:00

Abstract

Standardized estimates play an important role in model interpretation and the comparison of effect sizes. Although proper standardized estimation formulas for latent variable interaction effects have been available for over a decade and are used and cited both domestically and internationally, no systematic comparison of proper standardized estimates obtained from different estimation methods has been conducted to date. Through simulation experiments, the performance of standardized estimates for latent variable interaction effects was compared across different conditions among the product indicator method, Latent Moderated Structural Equations (LMS), Bayesian methods with uninformative priors, and Bayesian methods with informative priors. The results revealed that under normal conditions, LMS and informative Bayesian methods performed well; whereas under non-normal conditions, the product indicator method was relatively robust, but required a larger sample size (no less than 500). When the sample size was small and the correlation between exogenous latent variables was low, uninformative Bayesian methods could be used.

Full Text

Standardized Estimation for Latent Interaction Effects: Method Comparison and Selection Strategy

WEN Zhonglin, OUYANG Jinying, FANG Junyan
(School of Psychology & Center for Studies of Psychological Application, South China Normal University, Guangzhou 510631, China)

Abstract

Standardized estimation plays a crucial role in model interpretation and the comparison of effect sizes. Although the appropriate standardized estimation formula for latent interaction effects has been available for over a decade and is widely used and cited both domestically and internationally, no systematic comparison of appropriate standardized estimates obtained through different estimation methods has been conducted to date. Through simulation experiments, this study compares the performance of standardized estimates for latent interaction effects derived from the product indicator approach, Latent Moderated Structural Equations (LMS), and Bayesian methods with and without prior information across various conditions. The results indicate that under normal conditions, LMS and the informative Bayesian method perform well, whereas under non-normal conditions, the product indicator approach is relatively robust but requires a larger sample size (no less than 500). For small samples with low correlation among exogenous latent variables, the non-informative Bayesian method can be used.

Keywords: latent variable, interaction effect, product indicator approach, Latent Moderated Structural Equations, Bayesian method, standardized estimation
Classification Code: B841

In psychology and other social sciences, structural equation modeling can analyze relationships among multiple latent variables. Beyond linear relationships, interactions can reveal more complex relationships between variables. Consequently, analyzing latent interaction effects has become an important topic in both theoretical and empirical research.

Since parameters estimated from raw data (referred to as raw estimates) are influenced by the measurement units of variables, they cannot be directly compared. Researchers often convert raw parameter estimates into scale-invariant standardized estimates. In structural equation modeling, completely standardized estimation standardizes both indicators and latent variables, enabling comparison of estimated parameters (typically path coefficients or factor loadings). For latent interaction effect models, standardized estimation treats all latent variables in the model as Z-scores, which not only facilitates interpretation of interaction effects but also simplifies comparison of simple main effects.

For observed variables, SPSS can be used to analyze interaction effects, but its output of standardized estimates is inappropriate; instead, standardized variables (i.e., Z-scores) should be used for modeling and estimation (Wen et al., 2008). For latent variable models containing interaction effects, standardized estimation is more complex. For a long time, latent variable modeling software (such as Mplus) produced inappropriate standardized estimates for interaction effects. Wen et al. (2008) provided appropriate standardized estimation and proved its scale invariance (i.e., it does not change with measurement units, Wen et al., 2010), which is applicable in models with or without mean structures (Wu et al., 2011).

For standardized estimation of latent interaction effects, current statistical software outputs vary—some are appropriate, others are not, requiring careful distinction. When confusion is unlikely, "standardized estimation" can simply refer to appropriate standardized estimation.

Mplus 8.2 (Muthén & Muthén, 2019) and later versions provide appropriate standardized estimation for Latent Moderated Structural Equations (LMS) and Bayesian methods based on Wen et al. (2010)'s formulas (Asparouhov & Muthén, 2020; see also Mplus Technical Document No. 23), enabling users to obtain appropriate standardized estimates conveniently.

Comparing and selecting estimation methods constitutes an important component of theoretical research on latent interaction effects. To address issues of non-normality and nonlinearity introduced by interaction terms, researchers have proposed various modeling and estimation methods, including product indicator approaches, distribution analytic approaches (LMS is one such method), and Bayesian approaches, and have compared the performance of these methods under different conditions (Jackman et al., 2011; Kelava et al., 2011; Marsh et al., 2004; Marsh et al., 2012; Wen & Wu, 2010; Wen et al., 2013).

However, existing research has not systematically compared the standardized estimates of latent interaction effects obtained through different estimation methods. Since Mplus 8.2 and later versions provide appropriate standardized estimation, such comparisons have become feasible. Note that appropriate standardized estimates can be calculated using raw estimates (Wen et al., 2010; Wen et al., 2008), but raw estimates and their corresponding standard errors are not scale-invariant (Wu et al., 2014). Therefore, comparison results based on raw estimates may not generalize to appropriate standardized estimates. In particular, in the calculation formula for appropriate standardized estimation, some raw estimates appear in the numerator and some in the denominator. If the biases in the numerator and denominator are in the same direction and of similar magnitude, the bias in the appropriate standardized estimate could be small. Thus, despite existing literature comparing different estimation methods based on raw estimates of latent interaction effects, it remains necessary to compare their performance in terms of standardized estimates.

This paper assumes that the latent variables and their indicators discussed can reasonably be treated as continuous variables. First, we review the standardized model, appropriate standardized estimation, and its properties for latent interaction effects, and summarize the main estimation methods for latent interaction effects and comparative studies of these methods. Next, through a simulation experiment, we compare the performance of appropriate standardized estimates for latent interaction effects obtained by the product indicator approach, LMS, and Bayesian methods under various common conditions, and provide recommendations for selecting estimation methods.

2. Standardized Estimation for Latent Interaction Effect Models

Consider the structural equation model for latent interaction effects (Marsh et al., 2004):

$$\eta = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta \quad (1)$$

where ξ_1 and ξ_2 are exogenous latent variables, $\xi_1\xi_2$ is their interaction term, η is an endogenous latent variable, ζ is a residual term, and γ_1, γ_2, and γ_3 are path coefficients (as shown in Figure 1 [Figure 1: see original paper]). Since the intercept term and the means of ξ_1 and ξ_2 do not affect the estimation of interaction effects, they can be set to zero (Algina & Moulder, 2001; Marsh et al., 2004). The means of ξ_1, ξ_2, and ζ are zero, while the mean of $\xi_1\xi_2$ is typically not zero, so the mean of η is also typically not zero.

Figure 1. Path diagram of a latent interaction effect model containing one interaction term

If we follow the conventional approach and treat $\xi_1\xi_2$ as a variable (rather than the product of two variables) for standardization, the standardized form of equation (1) is:

$$\eta' = \gamma_1'\xi_1' + \gamma_2'\xi_2' + \gamma_3'(\xi_1\xi_2)' + \zeta' \quad (2)$$

where η' is the standardized variable of η, ξ_1' and ξ_2' are the standardized variables of ξ_1 and ξ_2, $(\xi_1\xi_2)'$ is the standardized variable of $\xi_1\xi_2$, ζ' is the residual term, and γ_1', γ_2', and γ_3' are conventional standardized coefficients. The relationship between conventional standardized estimates and raw estimates is (Wen et al., 2010; Wen et al., 2008):

$$\gamma_1' = \gamma_1\frac{sd(\xi_1)}{sd(\eta)}$$

$$\gamma_2' = \gamma_2\frac{sd(\xi_2)}{sd(\eta)}$$

$$\gamma_3' = \gamma_3\frac{sd(\xi_1\xi_2)}{sd(\eta)}$$

where $sd(\eta)$ denotes the raw estimate of the standard deviation of η, and similarly for the others.

However, equation (2) is not the standardized form we desire, because it standardizes $\xi_1\xi_2$ as a single variable, thus eliminating the interactive product term and making the results difficult to interpret in terms of interaction effects. The appropriate standardized form of equation (1) is:

$$\eta' = \alpha + \gamma_1'\xi_1' + \gamma_2'\xi_2' + \gamma_3''\xi_1'\xi_2' + \zeta'' \quad (4)$$

where α is the intercept term (since η' is a standardized variable while the means of ξ_1' and ξ_2' are typically not zero, the intercept is generally non-zero), γ_1', γ_2', and γ_3'' are appropriate standardized coefficients. Wen et al. (2008, see also Wen et al., 2010) derived the relationship between appropriate standardized estimates and conventional standardized estimates as follows:

$$\gamma_1'' = \gamma_1', \quad \gamma_2'' = \gamma_2', \quad \gamma_3'' = \gamma_3' \frac{sd(\xi_1)sd(\xi_2)}{sd(\xi_1\xi_2)}$$

where raw estimates of the standard deviations of ξ_1, ξ_2, and $\xi_1\xi_2$ are used. If appropriate standardized estimates are calculated directly using raw estimates (Mplus program see Appendix 1), the formulas are:

$$\gamma_1'' = \gamma_1 \frac{sd(\xi_1)}{sd(\eta)}, \quad \gamma_2'' = \gamma_2 \frac{sd(\xi_2)}{sd(\eta)}, \quad \gamma_3'' = \gamma_3 \frac{sd(\xi_1)sd(\xi_2)}{sd(\eta)}$$

The above appropriate standardized estimates are applicable to both latent and observed variables. Since the derivation does not involve specific estimation methods, they are applicable to different interaction effect estimation methods, such as product indicator approaches, distribution analytic approaches, and Bayesian methods. Moreover, appropriate standardized estimation is independent of variable distribution and does not require the assumption of normal distribution.

More importantly, appropriate standardized estimation has scale-free properties (Wen et al., 2010). Scale invariance has three meanings: First, standardized estimates of main effects and interaction effects are scale-invariant; second, standardized estimates of all factor loadings in measurement equations are scale-invariant; third, the standard errors and t-values corresponding to standardized estimates are also scale-invariant. Due to these properties, using data with different units yields identical standardized estimates.

The above appropriate standardized estimation has been applied in many studies on latent interaction effects. Numerous theoretical studies have employed appropriate standardized estimation in Monte Carlo simulations (e.g., Asparouhov & Muthén, 2020; Büchner & Klein, 2020; Foldnes & Hagtvet, 2014; Kelava et al., 2011; Wu et al., 2013). Brandt et al. (2015) extended appropriate standardized estimation to the mixture model framework. In personality psychology, developmental psychology, clinical psychology, and other fields, many applied studies cite Wen et al. (2010) and report standardized estimation results (e.g., Ito et al., 2015; Laczniak et al., 2017; Nagengast et al., 2011; You et al., 2016). The popular structural equation software Mplus, in version 8.2 or above, already supports LMS and Bayesian methods for analyzing latent interaction effect models and provides appropriate standardized estimation (Asparouhov & Muthén, 2020).

3.1 Product Indicator Approach

The product indicator approach was first proposed by Kenny and Judd (1984), using products of indicators as indicators of the latent interaction term, including constrained, partially constrained, and unconstrained models. The constrained and partially constrained models are complex and error-prone; only the unconstrained model proposed by Marsh et al. (2004) has been accepted by applied researchers (Wen & Liu, 2020). Therefore, unless otherwise specified, the product indicator approach refers to the unconstrained model hereafter.

When estimating product indicator models, Maximum Likelihood (ML) estimation is generally used. Although ML estimation requires indicators of latent variables to follow a multivariate normal distribution, it remains robust under non-normal conditions (Boomsma, 1983; Hau & Marsh, 2004), and standard errors and χ^2 can be corrected (Satorra & Bentler, 1994), which can be implemented in common SEM software including Mplus (see Appendix 1).

3.2 Latent Moderated Structural Equations

Distribution analytic approaches can directly address the non-normality issue of the interaction term $\xi_1\xi_2$ (Wen et al., 2013), including Latent Moderated Structural Equations (LMS, Klein & Moosbrugger, 2000; Schermelleh-Engel et al., 1998) and Quasi-Maximum Likelihood (QML, Klein & Muthén, 2007).

Currently, only LMS can be implemented in Mplus, while the QML method requires specialized software, so this paper only considers LMS. LMS approximates the joint distribution of indicators as a mixture of a finite number of conditional normal distributions and obtains maximum likelihood estimates of parameters through EM algorithm iterations, making it a generalized maximum likelihood method. Klein and Moosbrugger (2000) demonstrated that LMS is highly effective for estimating latent interaction effects. However, it is worth noting that while LMS accounts for the non-normality of the interaction term $\xi_1\xi_2$, it is still based on the assumption of normal distribution of latent variables, which can produce substantial bias when latent variables ξ_1 and ξ_2 are non-normally distributed (Cham et al., 2012; Kelava & Nagengast, 2012).

3.3 Bayesian Method

Introducing Bayesian methods in structural equation modeling can also be used to estimate latent interaction effects (Arminger & Muthén, 1998; Lee et al., 2007). Bayesian structural equation models combine prior distributions with likelihood functions given the data to determine posterior distributions of unknown parameters, using Markov chain Monte Carlo methods such as Gibbs sampling (Geman & Geman, 1984) and the Metropolis–Hastings algorithm (Hastings, 1970; Metropolis et al., 1953) to estimate posterior distributions. Bayesian methods provide credible estimates not only for simple structural equation models but also for nonlinear models (Lee & Song, 2004; Lee et al., 2007). In the past, Bayesian methods for latent interaction effect models were typically implemented using WinBugs software, but Mplus 8.2 and above can now implement them and provide appropriate standardized estimates.

Bayesian methods are often considered advantageous in situations with small samples and complex models such as nonlinear relationships among variables (Zhang et al., 2019). Since Bayesian methods combine prior information with observed data, as sample size decreases, the influence of prior information on posterior distributions increases while the influence of data decreases. Therefore, Bayesian methods do not require large-sample asymptotic theory and have inherent advantages in small-sample conditions (Gelman et al., 2014; Lee & Song, 2004; Lee et al., 2007; Muthén & Asparouhov, 2012). Lee et al. (2007) found in simulation studies that Bayesian methods produce acceptable estimates of latent interaction effects even with small samples (N = 150). However, the small-sample advantage of Bayesian methods is not absolute: without prior information, Bayesian methods may exhibit severe bias, performing worse than maximum likelihood estimation. Only when prior information is accurate can Bayesian methods perform acceptably in small samples (Smid et al., 2020). With very large sample sizes, Bayesian estimates are equivalent to maximum likelihood estimates (Gelman et al., 2014). Lee and Song (2004) found that Bayesian methods are not robust under non-normal conditions when estimating confirmatory factor analysis and structural equation models with small samples. Jia (2016) found that Bayesian structural equation models have excessively low true coverage rates for parameter confidence intervals under severely non-normal conditions.

3.4 Comparison of Estimation Methods for Latent Interaction Effects

For raw estimates, existing literature has compared the performance of product indicator approaches, LMS, and Bayesian methods in estimating latent interaction effects. Under normal or slightly skewed conditions, the unconstrained model shows acceptable estimation bias, standard error bias, and statistical power when sample size is large; LMS performs relatively better, with negligible estimation bias and standard error bias even in small samples, and higher statistical power (Cham et al., 2012; Jackman et al., 2011; Marsh et al., 2004). Under non-normal conditions, the unconstrained model shows acceptable estimation bias, standard error bias, coverage rates, and Type I error rates in large samples, while LMS exhibits bias in estimating nonlinear effects and their standard errors, with higher Type I error rates than the unconstrained model (Cham et al., 2012; Kelava & Nagengast, 2012; Marsh et al., 2004).

Currently, few studies systematically compare Bayesian methods with the other two methods. Lee et al. (2004) compared Bayesian methods with the constrained product indicator model and found that Bayesian methods perform well in small samples under normal latent variable distributions, while the constrained model is only acceptable with larger samples. Kelava and Nagengast (2012) systematically compared Bayesian methods within a mixture model framework with the unconstrained product indicator approach and LMS under different latent variable distributions, finding that Bayesian methods using two to four latent classes produce unbiased estimates of nonlinear effects under non-normal latent variable conditions, outperforming other methods. Asparouhov and Muthén (2020) compared LMS and non-informative Bayesian methods in a multilevel framework under normal, large-sample conditions (N = 1000), with results favoring Bayesian methods.

In summary, LMS outperforms the unconstrained product indicator approach under normal conditions; under non-normal conditions, LMS shows substantial bias, while the unconstrained product indicator approach shows acceptable estimation bias in large samples. Wen and Liu (2020) speculated that Bayesian methods might perform well under small-sample and non-normal conditions, but further research is needed.

However, since appropriate standardized estimates for latent interaction effects are calculated using formulas based on raw estimates (or together with conventional standardized estimates) rather than being direct estimates from a model, the above findings regarding the performance of estimation methods based on raw estimates may not be consistent with their performance based on appropriate standardized estimates. Whether consistent or not, specialized research is needed to draw conclusions; either way, the results would be meaningful. Below, we systematically compare the performance of appropriate standardized estimates for latent interaction effect models obtained by the product indicator approach, LMS, Bayesian methods with informative priors (BI), and Bayesian methods with non-informative priors (BN) under various common conditions, providing references for selecting estimation methods in applied research.

4.1 Simulation Study Design

This study examines the performance of appropriate standardized estimates for latent interaction effects obtained by the product indicator approach, LMS, and Bayesian methods under different conditions of exogenous latent variable distribution, correlation magnitude, and sample size. Following the simulation study design of Marsh et al. (2004), the data-generating structural model (true model) is:

$$\eta = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta$$

where the variance of the endogenous latent variable η is set to 1, and the two exogenous latent variables ξ_1 and ξ_2 are standardized variables (M = 0, SD = 1). Consequently, the main effects and interaction effect in the true model are standardized (unaffected by the mean of η).

Two distribution conditions for generating ξ_1 and ξ_2 are considered: normal and non-normal conditions. The non-normal distribution is obtained by weighting two $\chi^2(6)$ distributions (Marsh et al., 2004). The skewness and kurtosis of latent variables under both conditions are shown in Table 1 and Table 2 . The latent variables η, ξ_1, and ξ_2 each have three indicators: y_1, y_2, y_3; x_1, x_2, x_3; and x_4, x_5, x_6, respectively. The standardized factor loadings of indicators on their factors are all 0.7. Fixing the loadings is justified because previous research has found that loading magnitude essentially does not affect the comparison results of estimation methods (e.g., Marsh et al., 2004). All indicators are generated by adding random errors with the same distribution condition as the factors ξ_1 and ξ_2.

The two main effects γ_1 and γ_2 in the model are both fixed at 0.4, and the interaction effect γ_3 is set to 0 (for calculating Type I error rate) or 0.2 (for calculating statistical power). The correlation between exogenous latent variables ξ_1 and ξ_2 is set to 0, 0.3, or 0.7. When the interaction effect $\gamma_3 = 0.2$, for normal distributions, the effect size ρ^2 of the interaction term (i.e., the R-squared increment of the interaction term, Marsh et al., 2004) corresponding to different latent variable correlation magnitudes (0, 0.3, 0.7) are 4.00%, 4.36%, and 5.96%, respectively. For non-normal distributions, ρ^2 values are 4.00%, 5.07%, and 9.88%, respectively. Common effect sizes for interaction terms generally range between 3% and 8% (Champoux & Peters, 1987; Moulder & Algina, 2002), indicating that the research design is reasonable.

Sample sizes are set at N = 100, 200, and 500, which are typical sample sizes in applied research (Jaccard & Wan, 1995; Lee et al., 2007; MacCallum & Austin, 2000; Marsh et al., 2004).

In summary, the factorial design of the simulation study is as follows:

- Distribution of ξ_1 and ξ_2: normal distribution, non-normal distribution (between-subjects factor)
- Correlation between ξ_1 and ξ_2: $\phi_{12} = 0, 0.3, 0.7$ (between-subjects factor)
- Interaction effect: $\gamma_3 = 0, 0.2$ (between-subjects factor)
- Sample size: N = 100, 200, 500 (between-subjects factor)
- Estimation method: product indicator approach, LMS, Bayesian method without prior information, Bayesian method with prior information (between-subjects factor)

This is a $2 \times 3 \times 2 \times 3 \times 4$ design with 144 conditions, of which 36 are between-subjects conditions. For each between-subjects condition, 500 samples were generated using PRELIS 2.3 (Jöreskog & Sörbom, 1996) (i.e., 500 replications). The initial indicator variances in the simulated data were all 1, with factor loadings as specified above, and measurement errors had the same distribution as the generating conditions for ξ_1 and ξ_2. Following Marsh et al. (2004), the simulated data were linearly transformed to different measurement units:

$$y_1' = 1 + y_1/0.7, \quad y_2' = 7 + x_1/0.7, \quad x_2' = 6 + x_4/0.7, \quad x_5' = 2 + y_2,$$
$$y_3' = 6 + x_2, \quad x_3' = 9 + 3x_5, \quad x_6' = 3 + y_3; \quad x_1' = 8 + 2x_3; \quad x_4' = 7 + x_6;$$

This transformation yields raw data with intercept terms. When building the latent interaction effect model, all indicators are first centered. For each sample, Mplus 8.2 is used to estimate the latent interaction effect model, obtaining results from the product indicator approach, LMS, non-informative Bayesian method, and informative Bayesian method.

The product indicator approach uses a model without mean structure (Lin et al., 2010; Marsh et al., 2004; Wu et al., 2009), employs matched product indicators $x_{1x}4$, $x_{2x}5$, $x_{3x}6$ (Marsh et al., 2004), and uses the fixed loading method to specify measurement units. Regardless of the latent variable distribution, product indicators are not normally distributed. In Mplus, "ESTIMATOR=MLM" is specified to correct standard errors and model chi-square statistics using the Satorra-Bentler method (Satorra & Bentler, 1994). Based on Wen et al. (2008, see also Wen et al., 2010)' s formulas, the "Model Constraint" command in Mplus is used to transform raw estimates into appropriate standardized estimates (Mplus program see Appendix 1).

The appropriate standardized estimation for LMS is already built into Mplus 8.2 and above; users only need to request standardized solutions without recalculation. The Mplus program can be found in Wen and Liu (2020, pp. 193-194).

The Bayesian method uses Mplus default settings (Asparouhov & Muthén, 2020). Under informative prior conditions, prior distributions for factor loadings, main effects, and interaction effects are set to normal distributions with means equal to true values and variance of 0.01 (Muthén & Asparouhov, 2012). The prior distribution for the variance-covariance matrix of latent variables ξ_1 and ξ_2 is set to an inverse Wishart distribution with mean equal to true values, diagonal element variance set to 0.01. Since information about one element in the inverse Wishart distribution determines information about other elements (Gelman et al., 2014), off-diagonal element variances range between 0.50 and 0.74, with calculation formulas provided in Asparouhov and Muthén (2010). The appropriate standardized estimation for Bayesian methods is already built into Mplus 8.2 and above; users only need to request standardized solutions without recalculation. The Mplus program can be found in Wen and Liu (2020, pp. 199-200).

The simulation results reported below refer to appropriate standardized estimates. Results for the true interaction effect $\gamma_3 = 0$ are shown in Table 1 and Table 2, while results for $\gamma_3 = 0.2$ are shown in Table 3 and Table 4 .

4.2.1 Proper Solutions

A proper solution refers to a solution where the model converges and all parameter estimates are reasonable (e.g., variances and standard errors are non-negative). Calculations of parameter estimation bias, standard error bias, and other indicators are based only on proper solutions. As shown in Tables 1-4, the proportion of proper solutions for LMS and both Bayesian methods is close to or equal to 100% in all conditions. For the product indicator approach, when N = 100, the proportion of proper solutions gradually increases from $\phi_{12} = 0$

(average 69.6%), $\phi_{12} = 0.3$ (average 79.1%) to $\phi_{12} = 0.7$ (average 88.5%); when N = 200, the proportion exceeds 90% (except when latent variables are non-normal and $\phi_{12} = 0$); when N = 500, all proportions are close to or equal to 100%.

4.2.2 Bias of Standardized Estimates of Interaction Effects

When $\gamma_3 = 0$, the bias of the standardized estimate of the interaction effect equals the mean of the standardized estimate. A bias not exceeding 0.02 is considered acceptable (for the $\gamma_3 = 0.2$ case, a relative bias of 10% corresponds to a bias of 0.02). Under normal distribution of latent variables ξ_1 and ξ_2 (Table 1), the bias for both product indicator approach and LMS is less than 0.01; the bias for both Bayesian methods is less than 0.02. In summary, when the interaction effect does not exist, under normal conditions, the standardized estimates from all methods are very close to zero.

Under non-normal distribution of latent variables ξ_1 and ξ_2 (Table 2), the product indicator approach shows an estimation bias of 0.031 for the interaction effect when $\phi_{12} = 0$ and N = 100, and less than 0.01 in all other conditions. LMS and both Bayesian methods are not as robust as the product indicator approach under non-normal conditions; the bias of standardized estimates for the interaction effect increases with increasing ϕ_{12} (and thus increasing skewness and kurtosis of $\xi_1\xi_2$) and does not systematically decrease with increasing sample size. Except when $\phi_{12} = 0$, the bias exceeds 0.02. In terms of method comparison, under non-normal conditions and except when $\phi_{12} = 0$ and N = 100, the product indicator approach has the smallest bias, with the ranking generally being PI < BI < LMS < BN.

We focus on the $\gamma_3 = 0.2$ case (Tables 3 and 4). The relative bias of the standardized estimate of the interaction effect equals the value of the mean of the 500 sample standardized estimates minus the true value, divided by the true value. Generally, a relative bias (absolute value) less than 5% is considered ignorable (Hoogland & Boomsma, 1998), and a relative bias not exceeding 10% (Cham et al., 2012) or 15% (Bandalos, 2002; Muthén et al., 1987) is considered acceptable.

Under normal distribution of latent variables ξ_1 and ξ_2 (Table 3), for the product indicator approach, relative bias exceeds 10% when N = 100 and ($\phi_{12} = 0.3$ or 0.7), and when N = 200 and $\phi_{12} = 0.3$, and is less than 10% in all other conditions. The relative bias of LMS is close to or less than 5% in all conditions. The relative bias of the non-informative Bayesian method is less than 10% in all conditions. The relative bias of the informative Bayesian method is close to or less than 5% in all conditions. In terms of method comparison, when N = 100, the informative Bayesian method has the smallest relative bias, while the product indicator approach has the largest, with the ranking being BI < LMS, BN ≤ PI; when N = 200 or 500, LMS has the smallest relative bias, with the ranking generally being LMS < BI < BN < PI. In summary, under normal conditions, LMS and BI perform well, followed by BN, while the product indicator approach requires a larger sample (e.g., N = 500) to be acceptable.

Under non-normal distribution of latent variables ξ_1 and ξ_2 (Table 4), for the product indicator approach, the relative bias of the interaction effect exceeds 10% when N = 100 and ($\phi_{12} = 0$ or 0.3), and is close to or less than 10% in all other conditions. The robustness of LMS and both Bayesian methods is far inferior to that of the product indicator approach; the relative bias of standardized estimates for the interaction effect increases with increasing ϕ_{12} (and thus increasing skewness and kurtosis of $\xi_1\xi_2$) and does not systematically decrease with increasing sample size: when $\phi_{12} = 0$, relative bias is close to or less than 10%; when $\phi_{12} = 0.3$, it exceeds 10% (even exceeding 30%); when $\phi_{12} = 0.7$, it exceeds 30% (even exceeding 50%). In summary, when latent variables ξ_1 and ξ_2 are non-normally distributed, the larger the correlation between ξ_1 and ξ_2, the more severe the non-normality. As long as ξ_1 and ξ_2 have a substantial correlation, LMS and Bayesian methods should be avoided; only the product indicator approach can be used, but with a large sample size (e.g., N = 500).

4.2.3 Bias of Standard Error Estimates for Interaction Effects

In each experimental condition, the standard deviation of the 500 sample standardized estimates of the interaction effect (denoted as SD) is used as the true value of the standard error. The relative bias of the standard error of the interaction effect equals the mean of the standard errors (denoted as SE) minus the true value, divided by the true value. A relative bias (absolute value) within 10% is considered acceptable (Hoogland & Boomsma, 1998). Given that interaction terms in models are always non-normal, for the product indicator approach, standard errors are corrected using the Satorra-Bentler method (Satorra & Bentler, 1994). Results are shown in Tables 1-4.

First, consider the $\gamma_3 = 0$ case. Under normal distribution of latent variables ξ_1 and ξ_2 (Table 1), the relative bias of standard errors for the product indicator approach exceeds 10% when N = 100 or N = 200, and is less than 10% when N = 500; the relative bias for LMS and non-informative Bayesian methods is less than 10% in all conditions; except when N = 500 and $\phi_{12} = 0.7$, the relative bias for the informative Bayesian method exceeds 10%. In summary, the relative bias of standard error estimates for interaction effects is small and acceptable for LMS and BN, acceptable for the product indicator approach only with large samples, while BI consistently overestimates standard errors across conditions.

Under non-normal distribution of latent variables ξ_1 and ξ_2 (Table 2), the product indicator approach shows relative bias of standard errors exceeding 10% in 5 conditions with small samples or low correlation between ξ_1 and ξ_2 (i.e., N = 100 or $\phi_{12} = 0$), and less than 10% in the other 4 conditions; the relative bias for LMS and non-informative Bayesian methods is close to or less than 10% in all conditions; except when N = 500 and $\phi_{12} = 0.7$, the relative bias for the informative Bayesian method exceeds 10%. In summary, under non-normal conditions, the relative bias of standard error estimates for interaction effects is similar to that under normal conditions: LMS and BN show small and acceptable bias, while the product indicator approach and BI are acceptable only with large samples.

Now consider the $\gamma_3 = 0.2$ case. Under normal distribution of latent variables ξ_1 and ξ_2 (Table 3), the relative bias of standard errors for the four methods is similar to the $\gamma_3 = 0$ case and will not be repeated here.

Under non-normal distribution of latent variables ξ_1 and ξ_2 (Table 4), the product indicator approach shows relative bias of standard errors less than 10% when N = 100 and $\phi_{12} = 0$, when N = 200 and $\phi_{12} = 0.3$ (or 0.7), and when N = 500, and exceeds 10% in other conditions; LMS and non-informative Bayesian methods show relative bias close to or less than 10% in all conditions; the informative Bayesian method shows relative bias less than 10% when N = 200 and $\phi_{12} = 0.7$, and when N = 500, and close to or greater than 10% in other conditions. In summary, under non-normal conditions, the relative bias of standard error estimates for interaction effects is small and acceptable for LMS and BN, while those for the product indicator approach and BI are acceptable only with large samples.

4.2.4 Type I Error Rate and Statistical Power for Interaction Effect Testing

Since testing has scale invariance, test results for standardized estimates should be the same as or similar to those for raw estimates, so there is no need to emphasize testing of standardized estimates. Type I error rate refers to the probability of incorrectly rejecting the null hypothesis when it is true (here, $\gamma_3 = 0$), equal to the proportion of 500 samples that reject the null hypothesis at the 0.05 level. A Type I error rate below 0.075 is considered an acceptable standard (Bradley, 1978; Wu et al., 2013). Results are shown in Tables 1 and 2.

Under normal distribution of latent variables ξ_1 and ξ_2 (Table 1), the Type I error rate for the product indicator approach is close to or below 0.075 in all conditions; LMS is slightly above 0.075 in 3 correlation conditions when N = 100, and when N = 200 and $\phi_{12} = 0$, and below 0.075 in the other 5 conditions; the Type I error rate for the non-informative Bayesian method is below 0.075 in all conditions except when N = 200 and $\phi_{12} = 0$; the Type I error rate for the informative Bayesian method is below 0.075 in all conditions. In summary, LMS tends to have higher Type I error rates, which become acceptable with larger sample sizes; other methods have acceptable Type I error rates, with the informative Bayesian method having the lowest.

Under non-normal distribution of latent variables ξ_1 and ξ_2 (Table 2), the Type I error rate for the product indicator approach is below 0.075 in all conditions; LMS is below 0.075 only when N = 200 and $\phi_{12} = 0.3$, and above 0.075 in the other 8 conditions; the non-informative Bayesian method is below 0.075 when

N = 100 and ($\phi_{12} = 0$ or 0.3), and above 0.075 in the other 7 conditions; the informative Bayesian method is below 0.075 in 3 correlation conditions when N = 100, when N = 200 and $\phi_{12} = 0$ or 0.3, and when N = 500 and $\phi_{12} = 0$, and above 0.075 in the other 3 conditions. In summary, except for the product indicator approach, other methods have acceptable Type I error rates only with small samples or when the non-normality of the interaction term is not severe (i.e., low correlation between ξ_1 and ξ_2).

Statistical power of parameter testing refers to the probability of correctly rejecting the null hypothesis when it is false (here, $\gamma_3 = 0.2$), equal to the proportion of 500 samples that correctly reject the null hypothesis at the 0.05 level. Statistical power closer to 1 is better, with 80% typically desired. Results are shown in Tables 3 and 4.

Under normal distribution of latent variables ξ_1 and ξ_2 (Table 3), the statistical power for the product indicator approach is below 0.8 when N = 100 or 200, above 0.8 when N = 200 and $\phi_{12} = 0.7$, and below 0.8 under the other 5 conditions; the informative Bayesian method is close to or above 0.8 under all conditions. In summary, the statistical power for testing interaction effects is lowest for the product indicator approach, which requires larger samples to ensure power above 80%; BI has the highest power, with good power even when sample size is 100; LMS and BN have power between these two methods.

Under non-normal distribution of latent variables ξ_1 and ξ_2 (Table 4), statistical power results are similar to those under normal conditions. The statistical power of all methods increases with sample size and with the severity of non-normality of the interaction term (higher correlation between ξ_1 and ξ_2 indicates more severe non-normality).

When models contain interaction effects, conventional standardized estimates are inappropriate. Wen et al. (2008, see also Wen et al., 2010) provided appropriate standardized estimation for latent interaction effects and proved its scale invariance. This study systematically compares the performance of appropriate standardized estimates for latent interaction effects obtained by the product indicator approach, LMS, and both Bayesian methods under various common conditions of exogenous latent variable distribution, correlation magnitude, and sample size.

5.1 Simulation Study Results

Table 5 summarizes the simulation study results for standardized estimates of latent interaction effects obtained by different methods, alongside comparison results for raw estimates from different methods (Wen & Liu, 2020, pp. 202-203) for easy comparison.

The proportion of proper solutions for LMS and both Bayesian methods equals or approaches 100% under all conditions. The proportion of proper solutions for the product indicator approach increases with sample size and correlation among exogenous latent variables, being relatively low with small samples (N = 100) but approaching 100% with large samples (N = 500). This is similar to previous findings (e.g., Cham et al., 2012).

When exogenous latent variables are normally distributed, the performance of the product indicator approach and LMS in terms of estimation bias, standard error bias, Type I error rate, and statistical power for appropriate standardized estimates of interaction effects is similar to that for raw estimates (Cham et al., 2012; Jackman et al., 2011; Kelava et al., 2011; Marsh et al., 2004; Wen & Liu, 2020). Specifically, for the product indicator approach, estimation bias for interaction effects becomes ignorable with large samples, bias of standard errors corrected by the Satorra-Bentler method becomes acceptable with large samples, Type I error rates are acceptable in all conditions, and although statistical power is relatively low, it exceeds 80% with large samples. For LMS, estimation bias for interaction effects is ignorable under all conditions, standard error bias is acceptable, and both estimation bias and standard error bias are smaller than those of other methods with medium to large samples (N = 200 or above). LMS tends to have higher Type I error rates and significantly higher statistical power than the product indicator approach.

Without prior information, the bias of appropriate standardized estimates and standard errors for interaction effects obtained by Bayesian methods is acceptable under all conditions, Type I error rates are almost always acceptable, and statistical power increases with sample size and correlation among exogenous latent variables, exceeding 80% with large samples or high correlation. With correctly specified prior information, the bias of interaction effect estimates obtained by Bayesian methods is acceptable under all conditions and smaller than other methods with small samples; however, standard errors are consistently overestimated across conditions, which is not necessarily problematic as it indicates that actual fluctuations in parameter estimates are smaller than estimated; Type I error rates and statistical power are acceptable under all conditions and better than other methods.

Previous research has also found that informative Bayesian methods overestimate parameter standard errors. Holtmann et al. (2016) found that both correctly and incorrectly specified prior information overestimate parameter standard errors, with stronger prior information leading to greater standard error bias; however, Lee et al. (2007) found that correctly specified prior information yields more accurate standard error estimates than misspecified prior information, possibly due to different specifications of variance in prior distributions in their study.

When exogenous latent variables are non-normally distributed, the performance of the product indicator approach and LMS in terms of estimation bias, standard error bias, Type I error rate, and statistical power for appropriate standardized estimates is also similar to that for raw estimates (Cham et al., 2012; Kelava & Nagengast, 2012; Marsh et al., 2004; Wen & Liu, 2020). The product indicator approach shows acceptable estimation bias for interaction effects with large samples and is more robust than other methods; standard error bias is acceptable with large samples or high correlation among exogenous latent variables; Type I error rates are acceptable in all conditions and lower than other methods with large samples or high correlation; statistical power exceeds 80% with large samples. LMS shows acceptable estimation bias for interaction effects only when exogenous latent variables are uncorrelated ($\phi_{12} = 0$), with large bias under high correlation. LMS Type I error rates are almost always high, with correspondingly high statistical power. Notably, LMS estimation bias is not systematically affected by sample size when exogenous latent variables have medium or high correlation, consistent with Marsh et al. (2004)' s findings for another distribution analytic method, QML. Cham et al. (2012) also found that LMS is not robust under non-normal conditions. Unlike the present study, Cham et al. (2012) used non-normal distributions with higher kurtosis and skewness, resulting in even worse (less robust) LMS performance, while the product indicator approach remained robust under large-sample conditions.

Both Bayesian methods show acceptable estimation bias for interaction effects only when exogenous latent variables have low correlation ($\phi_{12} = 0$); with small samples and low correlation, both Bayesian methods show small estimation bias, but high bias under high correlation. Non-informative Bayesian methods have acceptable standard error bias, while informative Bayesian methods are acceptable only with large samples. Type I error rates and statistical power for both Bayesian methods increase with sample size or severity of non-normality of the interaction term (higher correlation between ξ_1 and ξ_2 indicates more severe non-normality).

Overall, the performance of standardized estimates for latent interaction effects obtained by various methods is very similar to that of raw estimates, indicating that there is no need to deliberately distinguish between calculating standardized or raw estimates when selecting estimation methods.

5.2 Performance of Bayesian Methods

Since Bayesian methods combine prior information with observed data, they are considered advantageous in small samples (Gelman et al., 2014; Lee & Song, 2004; Lee et al., 2007; Muthén & Asparouhov, 2012). However, research on Bayesian methods for latent interaction effect estimation is relatively scarce compared to other methods, especially regarding Type I error rates and statistical power, which remain largely unexplored (see Table 5).

This study finds that Bayesian methods are not necessarily more accurate than LMS in small-sample conditions; the non-informative Bayesian method is slightly worse than LMS in most conditions. Only when accurate prior information is added do Bayesian methods show certain advantages, but informative Bayesian methods overestimate standard errors. This is consistent with Smid et al. (2020)' s recent conclusion: without prior information, Bayesian methods may exhibit severe bias, performing worse than maximum likelihood estimation;

only with accurate prior information do Bayesian methods demonstrate excellent small-sample properties.

When prior information is accurate, the smaller the variance of the prior distribution, the more accurate the estimation results and the higher the statistical power (also see Fang et al., 2019; Miočević et al., 2017). However, if the prior information is incorrectly specified, the smaller the variance of the prior distribution, the less accurate the estimation results; only with large sample sizes can the problems caused by incorrect prior information be remedied (Fang et al., 2019; Lee et al., 2007). In actual research, prior information can be set based on previous studies or meta-analyses. If there is little confidence in these parameter values, lower precision can be set for the prior distribution, such as variance greater than 1 (Fang et al., 2019; Miočević et al., 2017). When prior information is difficult to obtain, other methods such as the product indicator approach or LMS can first be used to estimate the latent interaction effect model, and prior distributions can then be set based on the obtained estimates.

Some studies in other contexts have found that Bayesian methods are robust under non-normal conditions (including correlation among exogenous latent variables) (Gelman et al., 2014; Kelava et al., 2008; Kelava & Nagengast, 2012). However, our results show that Bayesian methods do not demonstrate advantages under non-normal conditions (especially with high correlation among exogenous latent variables). When exogenous latent variables are normal but highly correlated, although Bayesian methods perform better than the product indicator approach, they show no advantage over LMS; when exogenous latent variables are non-normally distributed, both Bayesian methods and LMS have high estimation bias for interaction effects, while the product indicator approach performs robustly. Other researchers have also found that Bayesian methods are not robust when estimating structural equation models under non-normal conditions (Jia, 2016; Lee & Song, 2004).

5.3 Recommendations for Applied Researchers

When variables are normally distributed, LMS is recommended for estimating latent interaction effects, as it provides the most accurate estimates of interaction effects and their standard errors with medium and large samples, and performs well even with small samples or high correlation among latent variables, though caution is needed regarding Type I error rates and effect sizes should be considered in inference. If accurate prior information can be obtained, Bayesian methods are preferable, especially in small-sample conditions, with small estimation bias, high statistical power, and acceptable Type I error rates, though standard errors are often overestimated and should be avoided for direct use. Non-informative Bayesian methods show no superiority over LMS. The product indicator approach is acceptable for estimating interaction effects and their standard errors with large samples, but has lower statistical power, requiring caution when interpreting non-significant results.

When variables are non-normally distributed, the unconstrained product indicator approach is recommended, but requires a large sample size, preferably no less than 500 (at which point estimation bias and standard error bias for interaction effects are acceptable, and statistical power and Type I error rates are acceptable). If the correlation between exogenous latent variables is low (which can be estimated and tested through confirmatory factor analysis), non-informative Bayesian methods can be considered for small samples, with all evaluation indicators being acceptable.

Notably, when using LMS and Bayesian methods, Mplus 8.2 or above will directly output appropriate standardized estimates for interaction effects when standardized solutions are requested; however, when using the unconstrained product indicator approach, either raw and conventional standardized solutions must first be obtained and then calculated using formula (5), or the program in Appendix 1 must be used (readers can easily adapt or modify it).

5.4 Limitations and Future Directions

First, in our study, each factor had an equal number of indicators, and all indicators had moderate loadings. Marsh et al. (2004)' s research indicates that loading magnitude has minimal impact on method comparison results. With unequal numbers of indicators, Wu et al. (2013) studied the performance of product indicator approach and LMS for latent interaction effect estimation, with results similar to this study, but the performance of Bayesian methods for appropriate standardized estimation and their comparison with LMS require further research.

Second, in some cases, Bayesian estimation bias first increases then decreases with increasing sample size. This phenomenon has been found in other Bayesian structural equation model estimations (Holtmann et al., 2016; Lee et al., 2007; van Erp et al., 2018; Yuan & MacKinnon, 2009). For informative prior cases, this may result from antagonistic effects of prior information and sample size on posterior distributions: with very small samples, accurate prior information can help estimate posterior distributions, while with large samples, sample information dominates (Lee et al., 2007). As sample size increases from small to moderate, the role of prior information weakens while sample information is not yet accurate enough, potentially making parameter estimates worse than with small or large samples. Non-informative Bayesian methods show similar phenomena, possibly indicating limited stability of Bayesian methods with moderate sample sizes, which requires deeper investigation.

In recent years, researchers have proposed new methods to address different problems in latent interaction effect models. Within the mixture model framework, some researchers have used Bayesian estimation (Kelava & Nagengast, 2012) or maximum likelihood estimation (Kelava et al., 2014) to analyze nonlinear latent variable models, with results showing these methods have smaller bias under non-normal conditions but are more complex than LMS (Asparouhov

& Muthén, 2020) and face the problem of appropriately selecting the number of latent classes (Kelava & Nagengast, 2012). Brandt et al. (2018) proposed a Bayesian Lasso method for estimating models with multiple interaction and quadratic effects, and Aytürk et al. (2020) used LMS to analyze latent interaction effect models with ordinal indicators. These new methods should consider appropriate standardized models and standardized estimation when establishing standardized models.

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: A Multidisciplinary Journal*, 8(1), 40–52.

Arminger, G., & Muthén, B. (1998). A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis-Hastings algorithm. *Psychometrika*, 63(3), 271–300.

Asparouhov, T., & Muthén, B. (2010). Bayesian analysis of latent variable models using Mplus. Muthén & Muthén.

Asparouhov, T., & Muthén, B. (2020). Bayesian estimation of single and multilevel models with latent variable interactions. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(2), 314–328.

Aytürk, E., Cham, H., Jennings, P. A., & Brown, J. L. (2020). Exploring the performance of latent moderated structural equations approach for ordered-categorical items. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(3), 410–422.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(1), 78–102.

Boomsma, A. (1983). *On the robustness of LISREL against small sample size and non-normality* [Unpublished doctoral dissertation]. University of Groningen.

Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31(2), 144–152.

Brandt, H., Cambria, J., & Kelava, A. (2018). An adaptive Bayesian lasso approach with spike-and-slab priors to identify multiple linear and nonlinear effects in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(6), 946–960.

Brandt, H., Umbach, N., & Kelava, A. (2015). The standardization of linear and nonlinear effects in direct and indirect applications of structural equation mixture models for normal and nonnormal data. *Frontiers in Psychology*, 6, 1813.

Büchner, R. D., & Klein, A. G. (2020). A Quasi-likelihood approach to assess model fit in quadratic and interaction SEM. *Multivariate Behavioral Research*, 55(6), 855–872.

Cham, H., West, S. G., Ma, Y., & Aiken, L. S. (2012). Estimating latent variable interactions with nonnormal observed data: A comparison of four approaches. *Multivariate Behavioral Research*, 47(6), 840–876.

Champoux, J. E., & Peters, W. S. (1987). Form, effect size and power in moderated regression analysis. *Journal of Occupational Psychology*, 60(3), 243–255.

Fang, J., Wen, Z., & Hau, K. -T. (2019). Mediation effects in 2-1-1 multilevel model: Evaluation of alternative estimation methods. *Structural Equation Modeling: A Multidisciplinary Journal*, 26(4), 591–606.

Foldnes, N., & Hagtvet, K. A. (2014). The choice of product indicators in latent variable interaction models: Post hoc analyses. *Psychological Methods*, 19(3), 444–457.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: CRC press.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6), 721–741.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), 97–109.

Hau, K. -T., & Marsh, H. W. (2004). The use of item parcels in structural equation modelling: Non-normal data and small sample sizes. *British Journal of Mathematical and Statistical Psychology*, 57(2), 327–351.

Holtmann, J., Koch, T., Lochner, K., & Eid, M. (2016). A comparison of ML, WLSMV, and Bayesian methods for multilevel structural equation models in small samples: A simulation study. *Multivariate Behavioral Research*, 51(5), 661–680.

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, 26(3), 329–367.

Ito, T. A., Friedman, N. P., Bartholow, B. D., Correll, J., Loersch, C., Altamirano, L. J., & Miyake, A. (2015). Toward a comprehensive understanding of executive cognitive function in implicit racial bias. *Journal of Personality and Social Psychology*, 108(2), 187–217.

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(2), 348–357.

Jackman, M. G. A., Leite, W. L., & Cochrane, D. J. (2011). Estimating latent variable interactions with the unconstrained approach: A comparison of methods to form product indicators for large, unequal numbers of items. *Structural Equation Modeling: A Multidisciplinary Journal*, 18(2), 274-288.

Jia, F. (2016). *Methods for handling missing non-normal data in structural equation modeling* [Unpublished doctoral dissertation]. University of Kansas.

Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User' s reference guide*. Chicago, IL: Scientific Software International, Inc.

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(2), 51-66.

Kelava, A., & Nagengast, B. (2012). A Bayesian model for the estimation of latent interaction and quadratic effects when latent variables are non-normally distributed. *Multivariate Behavioral Research*, 47(5), 717-742.

Kelava, A., Nagengast, B., & Brandt, H. (2014). A nonlinear structural equation mixture modeling approach for nonnormally distributed latent predictor variables. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(3), 468-481.

Kelava, A., Werner, C. S., Schermelleh-Engel, K., Moosbrugger, H., Zapf, D., Ma, Y., ···West, S. G. (2011). Advanced nonlinear latent variable modeling: Distribution analytic LMS and QML estimators of interaction and quadratic effects. *Structural Equation Modeling: A Multidisciplinary Journal*, 18(3), 465-491.

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, 96(1), 201-210.

Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, 65(4), 457-474.

Klein, A. G., & Muthén, B. O. (2007). Quasi-maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research*, 42(4), 647-673.

Laczniak, R. N., Carlson, L., Walker, D., & Brocato, E. D. (2017). Parental restrictive mediation and children' s violent video game play: The effectiveness of the Entertainment Software Rating Board (ESRB) rating system. *Journal of Public Policy & Marketing*, 36(1), 70-78.

Lee, S. Y., & Song, X. Y. (2004). Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research*, 39(4), 653-686.

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(1), 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: A Multidisciplinary Journal*, 14(3), 404-434.

Lin, G. C., Wen, Z., Marsh, H. W., & Lin, H. S. (2010). Structural equation models of latent interactions: Clarification of orthogonalizing and double-mean-centering strategies. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(3), 374-391.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, 51(1), 201-226.

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(3), 275-300.

Marsh, H. W., Wen, Z., Nagengast, B., & Hau, K. -T. (2012). Structural equation models of latent interaction. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 436-458). New York: Guilford Press.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6), 1087-1092.

Miočević, M., MacKinnon, D. P., & Levy, R. (2017). Power in Bayesian mediation analysis for small sample research. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(5), 666-683.

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

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, 17(3), 313-335.

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, 52(3), 431-462.

Muthén, L. K., & Muthén, B. O. (2019). *Mplus user' s guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K. -T., & Trautwein, U. (2011). Who took the "×" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science*, 22(8), 1058-1066.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.

Schermelleh-Engel, K., Klein, A., & Moosbrugger, H. (1998). Estimating nonlinear effects using a latent moderated structural equations approach. In R. E. Schumacker & G. A. Marcoulides (Eds.), *Interaction and nonlinear effects in structural equation modeling* (pp. 203–238). Mahwah, NJ: Erlbaum.

Smid, S. C., McNeish, D., Miočević, M., & van de Schoot, R. (2020). Bayesian versus frequentist estimation for structural equation models in small sample contexts: A systematic review. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(1), 131–161.

van Erp, S., Mulder, J., & Oberski, D. L. (2018). Prior sensitivity analysis in default Bayesian structural equation modeling. *Psychological Methods*, 23(2), 363–388.

Wen, Z., Hau, K. -T., & Marsh, H. W. (2008). Appropriate standardized estimates for moderating effects in structural equation models. *Acta Psychologica Sinica*, 40(6), 729–736.

Wen, Z., & Liu, H. (2020). *Analyses of mediating and moderating effects: Methods and applications*. Beijing: Educational Science Publishing House.

Wen, Z., Marsh, H. W., & Hau, K. -T. (2010). Structural equation models of latent interactions: An appropriate standardized solution and its scale-free properties. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(1), 1–22.

Wen, Z., & Wu, Y. (2010). Evolution and simplification of the approaches to estimating structural equation models with latent interaction. *Advance in Psychological Science*, 18(8), 1306–1313.

Wen, Z., Wu, Y., & Hau, K. -T. (2013). Latent interaction in structural equation modeling: Distribution-analytic approaches. *Psychological Exploration*, 33(5), 409–414.

Wu, Y., Wen, Z., Hau, K. -T., & Marsh, H. W. (2011). Appropriate standardized solutions in structural equation models with interaction effects without the mean structure. *Acta Psychologica Sinica*, 43(10), 1219–1228.

Wu, Y., Wen, Z., & Li, B. (2014). Testing the standardized solutions in structural equation models with interaction effects. *Psychological Exploration*, 34(3), 260–264.

Wu, Y., Wen, Z., & Lin, G. (2009). Structural equation modeling of latent interactions without using the mean structure. *Acta Psychologica Sinica*, 41(12), 1252–1259.

Wu, Y., Wen, Z., Marsh, H. W., & Hau, K. -T. (2013). A comparison of strategies for forming product indicators for unequal numbers of items in structural equation models of latent interactions. *Structural Equation Modeling: A Multidisciplinary Journal*, 20(4), 551-567.

You, J., Deng, B., Lin, M. P., & Leung, F. (2016). The interactive effects of impulsivity and negative emotions on adolescent nonsuicidal self-injury: A latent growth curve analysis. *Suicide and Life-Threatening Behavior*, 46(3), 266-283.

Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, 14(4), 301-322.

Zhang, L., Lu, J., Wei, X., & Pan, J. (2019). Bayesian structural equation modeling and its current researches. *Advance in Psychological Science*, 27(11), 1812-1825.

Appendix 1: Mplus Program for Unconstrained Product Indicator Approach to Latent Interaction Effects

```
TITLE: Unconstrained PI Approach
DATA: FILE is ex.dat;
VARIABLE: NAMES ARE y1-y3 x1-x6;
USEVARIABLES ARE x1-x6 y1-y3 x1x4 x2x5 x3x6;
DEFINE: x1x4 = x1*x4; x2x5 = x2*x5; x3x6 = x3*x6;  ! Define product indicators
ANALYSIS: ESTIMATOR = MLM; ! MLM command uses SB method to correct standard errors
MODEL: ! Specify measurement model
  eta by y1 y2 y3;
  ksi1 by x1 x2 x3;
  ksi2 by x4 x5 x6;
  ! Measurement model for latent product term
  ksi3 by x1x4 x2x5 x3x6;
  ! Specify structural model and label path coefficients
  eta on ksi1 (g1) ksi2 (g2) ksi3 (g3);  ! g1-g3 represent coefficients for ksi1-ksi3 in str
  ! Label variances, covariances, and residuals
  ksi1 (v1); ksi2 (v2); ksi3 (v3);   ! v1-v3 represent variances of ksi1-ksi3
  ksi1 with ksi2 (c12);   ! Covariance between ksi1 and ksi2
  ksi1 with ksi3 (c13);   ! Covariance between ksi1 and ksi3
  ksi2 with ksi3 (c23);   ! Covariance between ksi2 and ksi3
  eta (v4);  ! Residual variance of eta
MODEL CONSTRAINT:
  new(h1-h4);  ! Generate new parameters
  h4 = (g1**2)*v1+(g2**2)*v2+(g3**2)*v3+2*g1*g2*c12+2*g1*g3*c13+2*g2*g3*c23+v4; ! Variance o
  h1 = g1*((v1/h4)**(1/2));  ! Appropriate standardized estimate for main effect (coefficien
  h2 = g2*((v2/h4)**(1/2));  ! Appropriate standardized estimate for main effect (coefficien
  h3 = g3*((v1*v2/h4)**(1/2));  ! Appropriate standardized estimate for interaction effect (
```

Note: Figure translations are in progress. See original paper for figures.