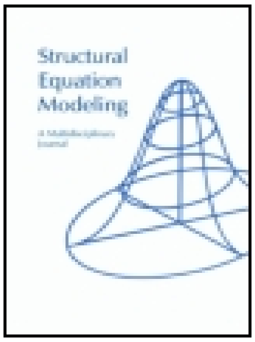

Routledge
Taylor & Francis Group

Structural Equation Modeling: A Multidisciplinary Journal

ISSN: 1070-5511 (Print) 1532-8007 (Online) Journal homepage: http://www.tandfonline.com/loi/hsem20

A Fit Index to Assess Model Fit and Detect Omitted Terms in Nonlinear SEM

Carla Gerhard, Rebecca D. Büchner, Andreas G. Klein & Karin Schermelleh-Engel

Structural Equation Modeling: A Multidisciplinary Journal, 24: 414–427, 2017

ISSN: 1070-5511 print / 1532-8007 online
DOI: 10.1080/10705511.2016.1268923

A Fit Index to Assess Model Fit and Detect Omitted Terms in Nonlinear SEM

Carla Gerhard, Rebecca D. Büchner, Andreas G. Klein, and Karin Schermelleh-Engel
Goethe University Frankfurt, Germany

A new descriptive fit measure, the Homoscedastic Fit Index (HFI), is proposed to detect omitted nonlinear terms (quadratic and interaction terms) in SEM by analyzing the dispersion of the residuals in the structural part of the model. The HFI is defined as a descriptive goodness-of-fit index for SEM. The Type I error rates of the HFI and the power to detect heteroscedasticity due to omitted nonlinear terms or nonnormally distributed variables are investigated in a Monte Carlo study. The results show that the new measure performs satisfactorily with regard to Type I error rates and power when sample size was sufficiently large. It is investigated under what conditions the Type I error rate was inflated. Nonnormally distributed error terms resulted in high power. Nonnormally distributed predictors had no influence on the Type I error rates.

Keywords: descriptive fit index, homoscedasticity, Monte Carlo study, nonlinear SEM

Over the last two decades, linear structural equation models have been extended to nonlinear structural equation modeling (SEM), by adding latent interaction or quadratic terms to the structural equation of the model (cf. Jöreskog & Yang, 1996; Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Marsh, Wen, & Hau, 2004; Ping, 1995; Schumacker & Marcoulides, 1998). Nonlinear effects, and in particular interaction and quadratic effects, are often relevant to psychological research. In differential psychology, for instance, nonlinear effects could be investigated to explain and predict behavior (cf. Dormann & Zapf, 2004). Nonlinear SEM is also common in social science research (cf. Beierlein, Werner, Preiser, & Wermuth, 2011; Berkel et al., 2010; Caravita, Di Blasio, & Salmivalli, 2009; Goodnight, Bates, Staples, Pettit, & Dodge, 2007; Lischetzke & Eid, 2003; Specht, Egloff, & Schmukle, 2011; Toker & Biron, 2012).

As with linear SEM, it is possible with nonlinear SEM to incorporate measurement errors in the model, to increase the construct validity by using multiple measures, and to estimate very complex linear and nonlinear relationships in a single model. For the analysis of nonlinear SEM, different estimation methods are available (for an overview see Klein & Muthén, 2007; Moosbrugger, Schermelleh-Engel, & Klein, 1997).

Frequently used methods are (a) the product indicator (PI) approaches (Algina & Moulder, 2001; Jaccard & Wan, 1995; Jöreskog & Yang, 1996; Kenny & Judd, 1984; Marsh et al., 2004; Ping, 1995, 1996; Wall & Amemiya, 2001), where products of the indicators are formed to be used as indicators of the latent nonlinear terms, and (b) the distribution analytic approaches of latent moderated structural equations (Klein & Moosbrugger, 2000) as well as quasi maximum likelihood (Klein & Muthén, 2007), which take the nonnormality caused by the latent nonlinear terms into account. Some newer developments are Bayesian approaches (e.g., Kelava & Nagengast, 2012; Lee, Song, & Tang, 2007; Song & Lu, 2010) and method of moment approaches (e.g., Mooijaart & Bentler, 2010; Wall & Amemiya, 2003).

Regardless of the method applied, a limitation of nonlinear SEM is that the evaluation of the model fit has not been thoroughly investigated compared to linear SEM. Whereas the overall model fit of linear SEM can be assessed using a likelihood ratio test or several descriptive fit measures (cf. Schermelleh-Engel, Moosbrugger, & Müller, 2003), established fit measures are not yet available for nonlinear SEM.

Corresponence should be addressed to Rebecca D. Büchner, Department of Psychology, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt am Main, Germany.
E-mail: buechner@psych.uni-frankfurt.de.

For linear SEM, the chi-square test is available. It is based on the likelihood ratio test (cf. Jöreskog, 1967). The likelihood of the target model is set in relation to the likelihood of a saturated model. Here, the saturated model is a model with zero degrees of freedom, which perfectly reproduces the empirical covariance matrix. The target model is nested within the saturated model and has less parameters than the number of variances and covariances given in the empirical covariance matrix.

For nonlinear models the situation is more difficult. The parameters of a nonlinear model are not fully identified when only variance and covariance information of the observed variables is used for estimation. The saturated model used for linear SEM cannot be applied to nonlinear SEM because the latter is not nested in the former model. Misspecifications due to omitted nonlinear terms cannot be readily detected by using the conventional chi-square test. Even though the chi-square values are not appropriate for nonlinear SEM, chi-square values are indeed calculated by the software when a PI approach is used. In a simulation study, Mooijaart and Satorra (2009) showed that the conventional chi-square test is not able to detect omitted interaction terms, and it has no power to identify this type of misspecification. Even when strong interaction effects are present and the model is misspecified, the distribution of the test statistic shows no deviation from a chi-square distribution. If a linear structural equation model fits the data well according to the chi-square statistic, the underlying model might indeed be severely nonlinear. The majority of the descriptive fit measures for linear SEM also build on the chi-square statistic (cf. Schermelleh-Engel et al., 2003), and possible nonlinearity in the model structure is not taken into account by these measures. However, to evaluate the fit of a nonlinear model, some researchers have recommended testing nonlinear models in a two-step estimation procedure (Maslowsky, Jager, & Hemken, 2015; Muthén, 2012): In a first step, a linear model with no nonlinear effects is tested to confirm a good fit in terms of the chi-square test. In a second step, the nonlinear terms are added, and the significance of a single or several nonlinear effects can be tested by the likelihood-ratio test statistic (Klein & Moosbrugger, 2000). A clear limitation of this approach is that the initial evaluation of the fit of the linear model could be misleading.

Other approaches to detect omitted nonlinear terms were developed in the context of regression analysis and SEM. The idea of these approaches is that omitted nonlinear terms result in heteroscedasticity.

Klein and Schermelleh-Engel (2010) developed the Z_{het} measure that quantifies the heteroscedasticity of the residual scores on a *Z*-scale. Z_{het} is based on the comparison of the likelihood for the structural residuals under the assumption of heteroscedasticity with the likelihood for the structural residuals under the assumption of homoscedasticity. A serious limitation of the Z_{het} measure is found in the need for a large sample size or very strong nonlinear effects to gain a desirable power of detection of heteroscedastic regression residuals (Klein, Gerhard, Büchner, Diestel, Schermelleh-Engel, 2016).

Klein et al. (2016) recently proposed a new measure for testing the heteroscedasticity of regression residuals, the h_{het} measure. The h_{het} measure functions by using the idea that, in the heteroscedastic case, regression residuals have been drawn from distributions with different standard deviations, thus the variance of the squared heteroscedastic residuals tends to be greater compared to homoscedastic residuals. The measure h_{het} is straightforward to apply. Its particular advantage is that it can detect the heteroscedasticity of the regression residuals that has been generated not only by observed but also by unobserved predictor variables.

In SEM, descriptive fit indexes are frequently used for the evaluation of model fit. Fit indexes, which are based on model comparisons, usually reach values between zero (indicating poor model fit) and one (indicating perfect model fit). As the exact distributions of these descriptive measures are often unknown, inferential statistical testing is not possible. Instead of using a statistical test, it is common to use cutoff values (Hu & Bentler, 1999), where good model fit is indicated when the descriptive fit values are greater than a certain threshold.[1] For example, the comparative fit index (CFI) has a cutoff value of .95 (Hu & Bentler, 1999).

The primary aim of this article is to develop a fit measure that has some descriptive validity, rather than a strict test of model fit. It is desirable to obtain a measure that is scaled simultaneously to other known fit indexes in SEM. The idea is to extend the measure h_{het} for residual scores taken from SEM and, thereby, to analyze the dispersion of the residuals in the structural part of the model. In SEM, the residual scores are not obtained directly, but they are score estimates for the latent residual variable. In particular, research in this article focused on the question whether the performance of the fit measure is influenced by the fact that now latent variable scores are used. Additionally, the fact that indicators do have a measurement error, as opposed to predicted scores in regression, might have an impact on how reliably the new fit measure performs.

This article is organized as follows: In the next section, we provide a novel measure for heteroscedasticity in SEM, the HFI. A Monte Carlo simulation study is then conducted to evaluate the performance of the HFI. Results for robustness, Type I error, and power rates are provided. Finally, we discuss implications of the simulation study and the limitations of the new measure.

[1] Examples of measures based on model comparisons are the nonnormed fit index (NNFI/TLI, Bentler & Bonett, 1980; Tucker & Lewis, 1973), the normed fit index (NFI, Bentler & Bonett, 1980), the comparative fit index (CFI, Bentler, 1990), the goodness-of-fit index (GFI, Jöreskog & Sörbom, 1989; Tanaka & Huba, 1984), and the adjusted goodness-of-fit index (AGFI, Jöreskog & Sörbom, 1989).

HOMOSCEDASTIC FIT INDEX

Omitted nonlinear terms in SEM cause heteroscedastic residuals in the structural part of the model. Therefore, a new descriptive index, the HFI, is provided to measure the homoscedasticity when a target structural equation model is fit. It is important to note that if the test indicates heteroscedasticity, different sources of heteroscedasticity are possible. In addition to omitted nonlinear terms, outliers in the data or misspecifications of the model equations are other potential sources of heteroscedasticity. The heteroscedasticity can also be affected by an unobserved nonlinear term (cf. Klein et al., 2016).

In the following, the measure of heteroscedasticity, h_{het}, for regression analysis is further developed and transformed to evaluate the heteroscedasticity of latent SEM. The measure h_{het} is defined as a standardized estimator of kurtosis:

$$h_{het} := \sqrt{\frac{n}{24}}\left(\frac{n^{-1}\sum e_i^4}{\left(n^{-1}\sum e_i^2\right)^2} - 3\right), \tag{1}$$

where e is the vector of residuals (Klein et al., 2016), and 3 and $\sqrt{n/24}$ standardize the expectation value and the standard deviation. Originally, this estimator is asymptotically standard normally distributed for independent random variables (cf. Davidson & MacKinnon, 1993). Also, because residuals meet the constraint $\bar{e} = 0$, they are not exactly independently distributed. In simulation studies, however, it was confirmed that the asymptotic distribution holds for sufficiently large sample sizes ($n \geq 100$). In addition, Klein et al. (2016) demonstrated that if the population errors are heteroscedastic and sample size is very large, h_{het} is asymptotically greater than zero. Thus, these results indicate the suitability of a one-tailed test for h_{het}.

For the HFI, the proposed model is compared to a homoscedastic comparison model: A close descriptive fit with a value close to one indicates that the residuals in the structural part of the model are homoscedastic. A value less than one indicates that the variance of the residuals of the target model is greater than the variance of a homoscedastic model, which means that the residuals are heteroscedastic. An advantage of a descriptive fit measure is that it quantifies the degree of fit along a continuum (cf. Hu & Bentler, 1999). It appears more appropriate to quantify the degree of heteroscedasticity along a continuum rather than by a simple positive or negative decision. Nevertheless, it is common to devise cutoff values for descriptive fit measures. The defined cutoff value might be treated as a more lenient decision criterion than the critical value of statistical significance test. As cutoff values between .90 and .97 are common for descriptive fit measures based on model comparisons in SEM (Hu & Bentler, 1999; Schermelleh-Engel et al., 2003), the HFI uses a cutoff value of .95.

For the development of the HFI we assume that residual scores $e_1, ..., e_n$ on a residual variable e are also available for the error ζ in the structural part of a structural equation model. When a structural equation model has been analyzed, such residual scores can be readily computed from the model equations by using the estimated parameters, the factor scores, and the observed scores on the indicators (for a more in-depth discussion, see Klein & Schermelleh-Engel, 2010).

Several steps are needed to transform h_{het} into the known terms of a descriptive fit measure for SEM. First, h_{het} is adjusted so that in the homoscedastic case the measure has a value close to 1, and in the heteroscedastic case its value is smaller than 1:

$$h_{het_{adj}} = \frac{1}{h_{het} + 1}. \tag{2}$$

In a second step, the required cutoff value for a descriptive measure is derived from Equation 2. The critical value of h_{het} in Equation 1 for a one-tailed test and for $\alpha = 5\%$ is $z = 1.645$. To determine a specific cutoff value, c_{HFI}, for $\alpha = 5\%$, the standardized value $z = 1.645$ and an additional scaling factor a are inserted in Equation 2:

$$c_{\text{HFI}} = \frac{1}{1.645a + 1}. \tag{3}$$

By solving Equation 3 for $c_{\text{HFI}} = .95$, an approximate value of $a = .032$ is obtained. A cutoff value of approximately .95 for the descriptive HFI is then given by:

$$c_{\text{HFI}} = \frac{1}{.032 \times 1.645 + 1} \approx .95. \tag{4}$$

As heteroscedasticity is given if $h_{het} > 0$, a one-tailed test is sufficient. For this, all values of $h_{het} \leq 0$ are set to 0 for the HFI. The descriptive measure HFI is then defined as:

$$\text{HFI} := \begin{cases} 1 & \text{if} \quad h_{het} \leq 0 \\ \frac{1}{.032 h_{het} + 1} & \text{if} \quad h_{het} > 0. \end{cases} \tag{5}$$

In the homoscedastic case, the HFI can be expected to assume a value close to 1.0, and in the heteroscedastic case, the HFI can be expected to assume a value smaller than 1.0. A cutoff value of $c_{\text{HFI}} = .95$ is recommended for $\alpha = 5\%$.

SIMULATION STUDY

In a Monte Carlo study we demonstrate the performance of the HFI for the detection of omitted nonlinear terms in structural equation models. The study investigates the influence of nonlinear effect size and sample size. Additionally, the effect of nonnormally distributed error terms is examined. Specifically, in a robustness study, the consequences of nonnormally distributed latent predictors on the HFI are investigated. Various linear and nonlinear population models were selected for data generation. After introducing the design of the simulation study, we present the findings on the performance of the HFI in detecting omitted nonlinear terms.

Population Models

Seven population models were chosen for data generation. The Type I error rate for overparameterized models was investigated using the first population model. The power and Type I error rate for models with omitted nonlinear terms were examined in the second, third, fourth, and fifth model. The influence of nonnormally distributed error terms was investigated in the sixth model. The seventh model focused on the Type I error rate when the assumption of normally distributed variables was violated.

The first population model, M_L, is linear and contains two linear effects:

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \zeta, \tag{6}$$

where η is a latent dependent variable, α is an intercept term, ξ_1 and ξ_2 are latent predictor variables, and ζ is an error term. The variables ξ_1, ξ_2, and ζ are normally distributed. Both of the latent predictors, ξ_1 and ξ_2, are measured by three indicators, x_1 to x_6, all of which have reliabilities of .64. The latent criterion η has zero mean and one indicator y with a reliability of 1.00. The correlation between ξ_1 and ξ_2 was set to $\Phi_{12} = .40$ in all conditions. The variances of ξ_1, ξ_2, and η were set to 1.0. The error terms correspond to 55.20% of the variance in η. The linear effect coefficients $\gamma_1 = .40$ and $\gamma_2 = .40$ were held constant across all simulation conditions.

The second model, M_{LQ}, is the same as the linear model M_L, except for the addition of the quadratic term ξ_1^2 to the structural equation. M_{LQ} is a quadratic model with two linear (L) effects and one quadratic effect (Q):

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2 + \zeta, \tag{7}$$

where the size of the quadratic effect ω_1 was set to .20, .25, and .30 in three effect size conditions, respectively. The quadratic effect explains between 8% and 18% of the variance in η so that 52% to 63% of the variance in the model is explained.

The third model, M_{LI}, is an interaction model with two linear (L) effects and one interaction effect (I):

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_3\xi_1\xi_2 + \zeta. \tag{8}$$

M_{LI} is the same as M_L, except for adding the interaction term $\xi_1\xi_2$. The size of the interaction effect ω_3 was set to .30, .35, and .40 in three effect size conditions. The interaction effect explains between 10% and 19% and the error term explains between 36% and 44% of the variance in η.

The fourth nonlinear equation model, M_{LQI}, contains two linear (L) effects, one quadratic (Q), and one interaction (I) term:

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2 + \omega_3\xi_1\xi_2 + \zeta. \tag{9}$$

The quadratic effect coefficient ω_1 and the interaction effect coefficient ω_3 were set to $\omega_1 = .20$, $\omega3 = .15$, to $\omega_1 = \omega_3 = .20$, and to $\omega_1 = .20$, $\omega_3 = .30$ in three effect size conditions, respectively. Taken together, these nonlinear effects explain between 10% and 19% of the variance in η. Overall, between 60% and 73% of variance in the model is explained. The structural part of M_{LQI} is depicted in Figure 1.

The fifth nonlinear equation model, M_{LQQI}, is the full model with two linear (L), two quadratic (QQ), and one interaction (I) term:

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2 + \omega_2\xi_2^2 + \omega_3\xi_1\xi_2 + \zeta. \tag{10}$$

The size of the quadratic effect coefficients ω_1 and ω_2 and the interaction effect coefficient ω_3 were set to $\omega_1 = \omega_2 = .10$, $\omega_3 = .15$, to $\omega_1 = \omega_2 = .10$, $\omega_3 = .20$, and to $\omega_1 = .15$, $\omega_2 = .10$, $\omega_3 = .20$ in three effect size conditions, respectively. These nonlinear effects together explain between 6% and 11.5% of the variance in η so that between 56% and 61% of variance in η is explained.

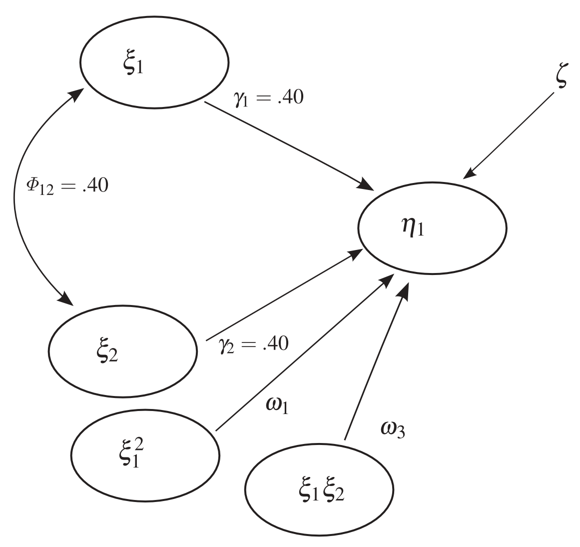

FIGURE 1 Structural part of the nonlinear population model, M_{LQI}, with two linear terms, ξ_1 and ξ_2, one quadratic term, ξ_1^2, and one interaction term, $\xi_1\xi_2$.

The sixth population model, $M_{S\zeta}$, is the same as the linear model M_L, except for the nonnormally distributed disturbance term ζ, $(S\zeta)$. The technical details for the nonnormally distributed disturbance terms are described in the Appendix. In the simulation study, the kurtosis of ζ was set to .80 with a skewness of .37, .49, or .70, and the skewness of ζ was set to .50 with a kurtosis of .38, .60, or 1.19 in six conditions. The error term corresponds to 55.20% of the variance in η.

The seventh population model, $M_{S\xi_2}$, is the same as the linear model M_L, except for a nonnormally distributed latent predictor $\xi_2, (S\xi_2)$. Derivation of nonnormally distributed latent predictors is provided in the Appendix. In four conditions, the values of kurtosis and skewness for ξ_2 were set to 0/0, .50/.30, .50/.50, and .95/.50. For this model, 44.80% of the variance is explained.

Design

The data for the population models were generated with the R software (*R* version 3.2.2, R Core Team, 2015). For each condition, 500 replications were performed, and the data were analyzed with M*plus* (Version 7.4: Muthén & Muthén, 2013) in the R software environment by using the M*plus* automation package (Hallquist & Wiley, 2015). The models were estimated with the latent moderated structural equations method (Klein & Moosbrugger, 2000). To obtain the residuals ζ_i for the structural part of the model in M*plus*, the factor scores option (i.e., 'Save = FS;') was used. This option provides ML-based factor scores for the latent variable η. The factor scores for η were subtracted from the scores for y. These values were studentized to obtain the required residual scores for ζ. This procedure has been described in detail by Klein and Schermelleh-Engel (2010). In addition, the authors provided the M*plus* syntax to obtain the residuals in the case of more than one indicator for η.

Across all conditions, the sample size n was set to 300, 500, 800, or 1,200. Data for seven population models (M_L, M_{LQ}, M_{LI}, M_{LQI}, M_{LQQI}, $M_{S\zeta}$, $M_{S\xi_2}$) were generated. The population model $M_{S\zeta}$ was analyzed as a linear model. The other population models M_L, M_{LQ}, M_{LI}, M_{LQI}, M_{LQQI}, and $M_{S\xi_2}$, were each analyzed as a correctly specified and as a misspecified model. For each model, several conditions were investigated.

Data generated for linear population models, M_L, were analyzed by correctly specified linear models, and by various overparameterized nonlinear models to study the Type I error rate. Data generated for nonlinear population models were analyzed by linear models for power analysis and by correctly specified models for Type I error analysis. Two linear models, $M_{S\zeta}$ and $M_{S\xi_2}$, were considered in the robustness study. Data produced for a linear population model with the nonnormally distributed error term, $M_{S\zeta}$, were each analyzed by linear models, M_L. For the investigation of linear models with a nonnormally distributed predictor term, $M_{S\xi_2}$, the generated data were analyzed by correctly specified linear models and underparameterized linear models where the nonnormally distributed predictor term was omitted. In the following, the percentage of data sets is reported for HFI values smaller than .95.

RESULTS

The HFI mean values, Type I error and power rates of the HFI to detect omitted nonlinear terms are reported.

The data for the linear population model M_L were analyzed by a correct linear model and by the overparameterized nonlinear models M_{LQ}, M_{LI}, M_{LQI}, and M_{LQQI}. In Table 1, mean HFI values and the Type I error rates of the HFI for the different linear and nonlinear analysis models are listed. The mean HFI values were .99 in all conditions. The HFI Type I error rates were close to the nominal $\alpha = 5\%$ level across all conditions. The HFI correctly indicates that the residuals are homoscedastic as no nonlinear terms have been omitted.

Table 2 presents the results for the quadratic population model M_{LQ}. The mean HFI values ranged between .98 and .99 when no term was omitted and between .82 and .98 when the quadratic term was not analyzed. The HFI Type I error rates were close to the nominal 5% level for $\omega_1 = .20$ and $\omega_1 = .25$, and they increased up to 9.40% for $\omega_1 = .30$ and $n = 1,200$. A desirable power of 80% was approximately reached in samples of $n = 600$ when the quadratic effect size was $\omega_1 = .30$ (not listed in Table 2). The power rates were close to 80% for $\omega_1 = .25$ and $n = 1,200$. No appropriate power was reached with a quadratic effect size of $\omega_1 = .20$.

Table 3 provides the results for the interaction population model M_{LI}. The mean HFI values were between .98 and .99 when no nonlinear term was omitted and between .85 and .97 when the interaction effect was not modeled. Most HFI Type I error rates were close to the nominal 5% level, and the largest value was 8.80%. The HFI reached a power of 80% in large samples ($n = 800$) when the interaction effect size was $\omega_3 = .40$.

Table 4 presents the results for a nonlinear population model M_{LQI} including a quadratic and an interaction effect. The mean HFI values were between .97 and .99 when the correctly specified models were analyzed and between .67 and .93 when the two nonlinear effects were not modeled. The Type I error rates were increased for an effect size of $\omega_1 = .20$ and $\omega_3 = .30$ and reached a maximum value of 21.20%. For the other effect sizes, the Type I error rates were slightly increased. A power of 80% was exceeded in samples with $n = 300$ when the nonlinear effect sizes were sufficiently large.

Table 5 provides the mean HFI values, the Type I error rates, and the power of the HFI for the full model M_{LQQI}. The Type I error rates for all correctly specified models were slightly higher with a maximum value of 12.80% and the mean HFI values were between .98 and .99. The mean HFI for omitted nonlinear

TABLE 1
Mean HFI Values and Type I Error Rates (in Percent) as a Function of Sample Size *(N)* for Linear Population Model M_L

Population Model	M_L $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2$									
Analysis Model	M_L $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2$		M_{LQ} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2$		M_{LI} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1\xi_2$		M_{LQI} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_3\xi_1\xi_2$		M_{LQQI} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2 + \omega_2\xi_2^2 + \omega_3\xi_1\xi_2$	
N	M	Type I Error	M	Type I Error	M	Type I Error	M	Type I Error	M	Type I Error
300	0.99	5.20	0.99	5.40	0.99	5.20	0.99	5.20	0.99	5.00
500	0.99	5.00	0.99	4.40	0.99	4.80	0.99	4.40	0.99	4.40
800	0.99	5.00	0.99	5.60	0.99	5.40	0.99	5.20	0.99	5.00
1,200	0.99	5.80	0.99	4.20	0.99	3.40	0.99	5.00	0.99	5.80

TABLE 2
Mean HFI Values, Type I Error Rates (in Percent), and Power (in Percent) as a Function of Sample Size *(N)* and Quadratic Effect Size for Population Model M_{LQ}

Population Model	M_{LQ} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2$											
Analysis Model	M_{LQ} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2$						M_L $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2$					
Nonlinear Effect	$\omega_1 = .20$		$\omega_1 = .25$		$\omega_1 = .30$		$\omega_1 = .20$		$\omega_1 = .25$		$\omega_1 = .30$	
N	M	Type I Error	M	Type I Error	M	Type I Error	M	Power	M	Power	M	Power
300	0.99	6.20	0.99	4.80	0.99	6.60	0.98	18.60	0.95	37.60	0.92	57.80
500	0.99	5.20	0.99	7.20	0.99	6.80	0.97	22.60	0.94	50.80	0.89	77.20
800	0.99	8.20	0.99	5.00	0.99	7.80	0.96	31.00	0.92	64.40	0.86	91.80
1,200	0.99	6.40	0.99	7.00	0.98	9.40	0.96	36.60	0.90	77.00	0.82	97.20

TABLE 3
Mean HFI Values, Type I Error Rates (in Percent), and Power (in Percent) as a Function of Sample Size *(N)* and Interaction Effect Size for Population Model M_{LI}

Population Model	M_{LI} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_3\xi_1\xi_2$											
Analysis Model	M_{LI} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_3\xi_1\xi_2$						M_L $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2$					
Nonlinear Effect	$\omega_3 = .30$		$\omega_3 = .35$		$\omega_3 = .40$		$\omega_3 = .30$		$\omega_3 = .35$		$\omega_3 = .40$	
N	M	Type I Error	M	Type I Error	M	Type I Error	M	Power	M	Power	M	Power
300	0.99	4.80	0.99	7.80	0.99	6.20	0.97	24.80	0.96	34.40	0.93	48.80
500	0.99	6.80	0.99	6.80	0.99	7.40	0.96	28.70	0.94	49.40	0.91	68.40
800	0.99	7.20	0.99	7.60	0.98	6.00	0.95	39.40	0.92	65.20	0.88	84.20
1,200	0.99	7.40	0.98	7.40	0.98	8.80	0.94	53.40	0.90	75.00	0.85	95.80

TABLE 4
Mean HFI Values Values, Type I Error Rates (in Percent), and Power (in Percent) as a Function of Sample Size *(N)* and Nonlinear Effect Size for Population Model M_{LQI}

Population Model	M_{LQI} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2 + \omega_3\xi_1\xi_2$											
Analysis Model	M_{LQI} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2 + \omega_3\xi_1\xi_2$						M_L $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2$					
Nonlinear Effect	$\omega_1 = .20;\ \omega_3 = .15$		$\omega_1 = \omega_3 = .20$		$\omega_1 = .20;\ \omega_3 = .30$		$\omega_1 = .20;\ \omega_3 = .15$		$\omega_1 = \omega_3 = .20$		$\omega_1 = .20;\ \omega_3 = .30$	
N	M	Type I Error	M	Type I Error	M	Type I Error	M	Power	M	Power	M	Power
300	0.99	6.80	0.99	7.40	0.98	11.40	0.93	49.60	0.91	66.00	0.83	91.80
500	0.99	8.20	0.98	9.40	0.98	14.40	0.91	64.00	0.88	81.80	0.78	98.20
800	0.98	8.60	0.98	9.20	0.97	17.60	0.88	83.00	0.84	94.20	0.72	99.80
1,200	0.99	6.60	0.98	12.60	0.97	21.20	0.86	89.40	0.82	98.00	0.67	100.00

TABLE 5
Mean HFI Value, Type I Error Rates (in Percent), and Power (in Percent) as a Function of Sample Size *(N)* and Nonlinear Effect Size for Population Model M_{LQQI}

Population Model	M_{LQQI} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2 + \omega_2\xi_2^2 + \omega_3\xi_1\xi_2$											
Analysis Model	M_{LQQI} $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_1\xi_1^2 + \omega_2\xi_2^2 + \omega_3\xi_1\xi_2$						M_L $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2$					
Nonlinear Effect	$\omega_1 = \omega_2 = .10;\ \omega_3 = .15$		$\omega_1 = \omega_2 = .10;\ \omega_3 = .20$		$\omega_1 = .15;\ \omega_2 = .10$		$\omega_1 = \omega_2 = .10;\ \omega_3 = .15$		$\omega_1 = \omega_2 = .10;\ \omega_3 = .20$		$\omega_1 = .15;\ \omega_2 = .10;\ \omega_3 = 20$	
N	M	Type I Error	M	Type I Error	M	Type I Error	M	Power	M	Power	M	Power
300	0.99	6.20	0.99	6.40	0.98	7.60	0.95	33.00	0.93	50.60	0.90	67.00
500	0.99	6.60	0.99	7.40	0.98	9.40	0.94	45.00	0.91	66.40	0.86	84.20
800	0.99	9.00	0.98	10.00	0.98	12.80	0.92	59.80	0.88	81.60	0.83	94.60
1,200	0.98	8.20	0.98	10.20	0.98	11.80	0.91	74.40	0.86	93.20	0.79	99.00

TABLE 6
Mean HFI Values and Type I Error Rates (in Percent) as a Function of Sample Size *(N)*, Kurtosis, and Skewness of ζ for Population Model $M_{S\zeta}$ with Nonnormally Distributed ζ

Population Model	$M_{S\zeta}$ $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2$											
Analysis Model	M_L $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2$											
(Centered) Kurtosis of ζ	0.80		0.80		0.80		0.38		0.60		1.19	
Skewness of ζ	0.37		0.49		0.70		0.50		0.50		0.50	
N	M	Power	M	Power	M	Power	M	Power	M	Power	M	Power
300	0.95	40.00	0.94	44.80	0.94	44.80	0.97	21.60	0.95	36.60	0.93	53.60
500	0.93	53.60	0.92	62.40	0.93	60.00	0.96	29.20	0.94	50.80	0.90	70.40
800	0.91	67.40	0.90	78.40	0.91	76.60	0.95	42.20	0.93	64.00	0.87	85.00
1,200	0.89	80.80	0.89	86.20	0.88	89.60	0.94	54.80	0.91	74.80	0.85	94.20

TABLE 7
Mean HFI Values, Type I Error Rates (in Percent), and Power (in Percent) as a Function of Sample Size *(N)* and Skewness of ξ_2 for Population Model $M_{S\xi_2}$

Population Model	$M_{S\xi_2}$ $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2$															
Analysis Model	M_L $\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2$								M_L $\eta = \alpha + \gamma_1\xi_1$							
(Centered) Kurtosis of ξ_2	0.00		0.50		0.50		0.95		0.00		0.50		0.50		0.95	
Skewness of ξ_2	0.00		0.30		0.50		0.50		0.00		0.30		0.50		0.50	
N	M	Type I Error	M	Type I Error	M	Type I Error	M	Type I Error	M	Power	M	Power	M	Power	M	Power
300	0.99	5.20	0.99	5.20	0.99	5.00	0.99	4.00	0.99	3.40	0.97	17.20	0.98	14.20	0.96	32.20
500	0.99	4.40	0.99	4.00	0.99	3.00	0.99	3.80	0.99	3.80	0.97	21.40	0.97	21.60	0.95	40.00
800	0.99	5.80	0.99	4.40	0.99	4.00	0.99	4.60	0.99	4.80	0.96	31.20	0.97	26.40	0.94	52.20
1,200	0.99	6.00	0.99	5.20	0.99	3.20	0.99	4.80	0.99	4.60	0.96	39.00	0.96	35.80	0.93	62.80

terms was between .79 and .95. A power of 80% was reached for $\omega_1 = \omega_2 = .10$, $\omega_3 = .20$ and $n = 800$, and for $\omega_1 = .15$, $\omega_2 = .10$, $\omega_3 = .20$ and $n = 500$.

The linear model with nonnormally distributed error terms, $M_{S\zeta}$, was analyzed as a linear model (Table 6). The power reached values between 40% and 95%, whereas both increasing kurtosis and skewness resulted in higher power. The mean HFI values were not greater than .95.

The linear population model $M_{S\xi_2}$ with skewed ξ_2 was analyzed as a linear model with one or two predictors (Table 7). For the analysis with two predictors, the Type I error rates ranged between 3% and 6% and the mean HFI values were .99. Power rates for the analysis with one predictor were also close to 5% for skewness and kurtosis with values close to 0. With increasing kurtosis and skewness, the power increased to between 17.20% and 62.80%, respectively.

DISCUSSION

In this article we proposed a novel descriptive measure for nonlinear structural equation models to detect omitted nonlinear terms, the HFI. The HFI examines the homoscedasticity of a target latent variable model. The HFI measure makes direct use of the dispersion of the squared residuals of the structural part of the model to examine a possible deviation from homoscedasticity. The HFI is proposed for the evaluation of fit, it quantifies if there is a sufficient modeling of possible nonlinearity in the model. Values close to one indicate that there is sufficient modeling of nonlinearity.

In a Monte Carlo study we demonstrated that the HFI is sensitive to heteroscedasticity caused by omitted nonlinear terms. The HFI responds to separately omitted quadratic and interaction terms, as well as to both terms omitted simultaneously. In sufficiently large sample and effect size conditions, the power analysis showed good results. Depending on the expected nonlinear effect sizes, a sample size greater than 500 is recommended for detecting heteroscedasticity in nonlinear SEM. In sufficiently large samples, the HFI appears to be a suitable method for the detection of omitted nonlinear terms in SEM.

The Type I error analysis showed mostly satisfactory results, keeping the 5% nominal level in most conditions for nonlinear effects that are not too strong. When analyzing linear population models, the 5% error rate was kept in all conditions. In some conditions, when the target model had large or multiple nonlinear terms, Type I error rates were inflated. Therefore, we recommend choosing more conservative cutoff values for models that already include some nonlinear relationships. For the models tested here, as a rule of thumb, results suggest reducing the recommended cutoff value of $c_{HFI} = .95$ by .01 for each significant nonlinear effect already included in the structural equation model. For example, when testing a model with a significant quadratic effect for possibly omitted nonlinear terms, a cutoff value of $c_{HFI} = .94$ seems to work satisfactorily.

Moreover, we investigated the influence of nonnormal latent predictor variables in a robustness study with satisfactory results. Findings revealed that the nonnormally distributed latent predictors did not influence the Type I error rates. The power to detect omitted nonnormally distributed predictor terms depended on the values of the kurtosis: An omitted normally distributed latent predictor (with kurtosis and skewness of zero) did not influence the test, resulting in a correct nominal Type I error level of almost 5%. Omitted nonnormally distributed latent predictors, comparable to nonlinear terms with large kurtosis and skewness values, clearly resulted in an increased power.

In the simulation study, nonnormally distributed error terms resulted in high power rates, which reflects the fact that the measure might also respond to other sources of heterogeneity beside nonlinearity.

Therefore, a two-step approach is recommended for researchers: In a first step, the researcher can assess the homoscedasticity of a latent model with the HFI measure. When the residuals are homoscedastic, the HFI value is close to one and it can be assumed that no strong nonlinear relationship has been overlooked. If the HFI indicates heteroscedastic residuals, the second step entails searching for the source of heteroscedasticity. If particular nonlinear relationships are assumed, they can be tested by the likelihood-based model difference test (T_D) for latent nonlinear effects (Klein & Moosbrugger, 2000). The use of the two robust statistics developed for conditions of nonnormality and small samples, the robust T_{DR} and the strictly positive T_{DRP} (Satorra & Bentler, 2001, 2010), might not be advisable for testing nonlinear effects, as they produce many negative difference values, low power, or both (Gerhard et al., 2015). If T_D shows a significant result, the nonlinear terms should be included in the target model. Future research should examine if the HFI can be used for the more common application context of detecting heteroscedastic residuals in general.

Although the HFI has been shown to detect omitted nonlinear terms in SEM, this study is limited to the conditions examined here. To further strengthen the position of the HFI within the SEM framework, future research should investigate the performance of the HFI in a number of additional conditions, such as higher model complexity or testing models with more than a single latent dependent variable. Moreover, the suitability of the HFI for SEM using PI approaches needs further testing as well. Future research could also investigate variants of the HFI measure for models with multiple indicators of the latent dependent variables. Overall, based on the simulation findings presented here, we believe that the HFI offers a promising new method for determining the heteroscedasticity of the residuals and for the detection of omitted nonlinear terms in SEM.

FUNDING

This research was supported by Grant No. SCHE1412-1/1 from the German Research Foundation (DFG).

ORCID

Rebecca D. Büchner http://orcid.org/0000-0002-1921-803X
Karin Schermelleh-Engel http://orcid.org/0000-0002-6341-0803

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. doi:10.1207/S15328007SEM0801_3

Beierlein, C., Werner, C. S., Preiser, S., & Wermuth, S. (2011). Are just-world beliefs compatible with justifying inequality? Collective political efficacy as a moderator. *Social Justice Research*, *24*, 278–296. doi:10.1007/s11211-011-0139-2

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246. doi:10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606. doi:10.1037/0033-2909.88.3.588

Berkel, C., Knight, G. P., Zeiders, K. H., Tein, J.-Y., Roosa, M. W., Gonzales, N. A., & Saenz, D. (2010). Discrimination and adjustment for Mexican American adolescents: A prospective examination of the benefits of culturally related values. *Journal of Research on Adolescence* *20*, 893–915. doi:10.1111/j.1532-7795.2010.00668.x

Caravita, S. C. S., Di Blasio, P., & Salmivalli, C. (2009). Unique and interactive effects of empathy and social status on involvement in bullying. *Social Development*, *18*, 140–163. doi:10.1111/j.1467-9507.2008.00465.x

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. New York NY: Oxford University Press.

Dormann, C., & Zapf, D. (2004). Customer-related social stressors and burnout. *Journal of Occupational Health Psychology* *9*(1), 61–82. doi:10.1037/1076-8998.9.1.61

Gerhard, C., Klein, A. G., Schermelleh-Engel, K., Moosbrugger, H., Gäde, J., & Brandt, H. (2015). On the performance of likelihood-based difference tests in nonlinear structural equation models. *Structural Equation Modeling*, *22*, 276–287. doi:10.1080/10705511.2014.935752

Goodnight, J. A., Bates, J. E., Staples, A. D., Pettit, G. S., & Dodge, K. A. (2007). Temperamental resistance to control increases the association between sleep problems and externalizing behavior development. *Journal of Family Psychology*, *21*(1), 39–48. doi:10.1037/0893-3200.21.1.39

Hallquist, M., & Wiley, J. F. (2015). Automating Mplus model estimation and interpretation (version 3) [computer software]. Vienna Austria: R Foundation for Statistical Computing. Retrieved from https://cran.r-project.org/web/packages/MplusAutomation/MplusAutomation.pdf

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. doi:10.1080/10705519909540118

Jaccard, J., & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin*, *117*, 348–357. doi: 10.1037/0033-2909.117.2.348

Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika* *32*, 443–482. doi:10.1007/BF02289658

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

Jöreskog, K. G., & Yang, F. (1996). Nonlinear structural equation models: The Kenny–Judd model with interaction effects. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 57–88). Mahwah, NJ: Erlbaum.

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*, 717–742. doi:10.1080/00273171.2012.715560

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, *96*(1), 201– 210. doi:10.1037/0033-2909.96.1.201

Klein, A. G., Gerhard, C., Büchner, R. D., Diestel, S., & Schermelleh-Engel, K. (2016). The detection of heteroscedasticity in regression models for psychological data. *Psychological Test and Assessment Modeling*, *58*, 567–592.

Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*, 457–474. doi:10.1007/BF02296338

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*, 647–673. doi:10.1080/00273170701710205

Klein, A. G., & Schermelleh-Engel, K. (2010). Introduction of a new measure for detecting poor fit due to omitted nonlinear terms in SEM. *AStA Advances in Statistical Analysis*, *94*, 157–166. doi:10.1007/s10182-010-0130-5

Lee, S.-Y., Song, X.-Y., & Tang, N.-S. (2007). Bayesian methods for analyzing structural equation models with covariates, interaction, and quadratic latent variables. *Structural Equation Modeling*, *14*, 404–434. doi:10.1080/10705510701301511

Lischetzke, T., & Eid, M. (2003). Is attention to feelings beneficial or detrimental to affective well-being? Mood regulation as a moderator variable. *Emotion*, *3*, 361–377. doi:10.1037/1528-3542.3.4.361

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. doi:10.1037/1082-989X.9.3.275

Maslowsky, J., Jager, J., & Hemken, D. (2015). Estimating and interpreting latent variable interactions: A tutorial for applying the latent moderated structural equations method. *International Journal of Behavioral Development*, *39*(1), 87–96. doi:10.1177/0165025414552301

Mooijaart, A., & Bentler, P. M. (2010). An alternative approach for non-linear latent variable models. *Structural Equation Modeling*, *17*, 357–373. doi:10.1080/10705511.2010.488997

Mooijaart, A., & Satorra, A. (2009). On insensitivity of the chi-square model test to nonlinear misspecification in structural equation models. *Psychometrika*, *74*, 443–455. doi:10.1007/s11336-009-9112-5

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

Muthén, B. O. (2012, September). *Latent variable interactions*. Retrieved from https://www.statmodel.com/download/LV%20Interaction.pdf

Muthén, L. K., & Muthén, B. O. (2013). M*plus* (version 7.1) [Computer software]. Los Angeles, CA: Muthén & Muthén.

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

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

R Core Team. (2015). *R: A language and environment for statistical computing* (Version 3.2.2). Vienna, Austria: R Foundation for Statistical Computing.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. doi:10.1007/BF02296192

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, *75*, 243–248. doi:10.1007/s11336-009-9135-y

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

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

Song, X.-Y., & Lu, Z.-H. (2010). Semiparametric latent variable models with Bayesian p-splines. *Journal of Computational and Graphical Statistics*, *19*(3), 590–608. doi:10.1198/jcgs.2010.09094

Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology*, *101*, 862–882. doi:10.1037/a0024950

Tanaka, J. S., & Huba, G. J. (1984). Confirmatory hierarchical factor analyses of psychological distress measures. *Journal of Personality and Social Psychology*, *46*, 621–635. doi:10.1037/0022-3514.46.3.621

Toker, S., & Biron, M. (2012). Job burnout and depression: Unraveling their temporal relationship and considering the role of physical activity. *Journal of Applied Psychology*, *97*, 699–710. doi:10.1037/a0026914

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1–10. doi:10.1007/BF02291170

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

Wall, M. M., & Amemiya, Y. (2003). A method of moments technique for fitting interaction effects in structural equation models. *British Journal of Mathematical and Statistical Psychology*, *56*, 47–63. doi:10.1348/000711003321645331

APPENDIX
NONNORMAL DISTRIBUTIONS

Nonnormal Error Term

The nonnormally distributed structural error ζ was generated as a mixture of $X \sim \mathcal{N}(0,1)$ and $Z \sim \chi^2_{df}$ with *df* degrees of freedom:

$$\zeta = \sigma^2_\zeta \left(c_1 X + \sqrt{1 - c_1^2} \frac{Z - df}{\sqrt{2df}} \right), \tag{A.1}$$

where $c1$ is a weight parameter and $\mathrm{Var}(\zeta) = \sigma^2_\zeta$. The (centered) kurtosis and the skewness can be expressed as

$$\mathrm{Kurt}(\zeta) = \frac{12}{df}(1 - c_1^2)^2 \tag{A.2}$$

$$\mathrm{Skew}(\zeta) = \frac{4}{\sqrt{2df}} \left(\sqrt{1 - c_1^2} \right)^3. \tag{A.3}$$

Nonnormal Predictor

The nonnormal predictor was generated as a mixture:

$$\xi_2 = .40\xi_1 + \sqrt{1 - .40^2} \left(c_2 X + \sqrt{1 - c_2^2} \frac{Z - df}{\sqrt{2df}} \right), \tag{A.4}$$

where $X \sim \mathcal{N}(0,1)$ and $Z \sim \chi^2 df$ with *df* degrees of freedom. Under this condition, the correlation between ξ_1 and ξ_2 is $\Phi_{12} = .40$ and the variance of ξ_2 is 1. The formula for the (centered) kurtosis of ξ_2 is

$$\mathrm{Kurt}(\xi_2) = (1 - .40^2)^2 (1 - c_2^2)^2 \frac{12}{df} - 3. \tag{A.5}$$

The formula for the skewness can be expressed as

$$\mathrm{Skew}(\xi_2) = \left(1 - .40^2\right)^{\frac{3}{2}} \left(1 - c_2^2\right)^{\frac{3}{2}} \frac{4}{\sqrt{2df}}. \tag{A.6}$$