Psicothema

Psicothema 2020, Vol. 32, No. 1, 115-121
doi: 10.7334/psicothema2019.235

ISSN 0214 - 9915 CODEN PSOTEG

www.psicothema.com

Specification issues in nonlinear SEM: The moderation that wasn't

Karina Rdz-Navarro[1] and Fan Yang-Wallentin[2]
[1] Universidad de Chile and [2] University of Uppsala

Abstract

Background: Analysis of interaction or moderation effects between latent variables is a common requirement in the social sciences. However, when predictors are correlated, interaction and quadratic effects become more alike, making them difficult to distinguish. As a result, when data are drawn from a quadratic population model and the analysis model specifies interactions only, misleading results may be obtained. **Method:** This article addresses the consequences of different types of specification error in nonlinear structural equation models using a Monte Carlo study. **Results:** Results show that fitting a model with interactions when quadratic effects are present in the population will almost certainly lead to erroneous detection of moderation effects, and that the same is true in the opposite scenario. Simultaneous estimation of interactions and quadratic effects yields correct results. **Conclusions:** Simultaneous estimation of interaction and quadratic effects prevents detection of spurious or misleading nonlinear effects. Results are discussed and recommendations are offered to applied researchers.

Keywords: Nonlinear structural equations, moderation, interaction effects, quadratic effects, model specification.

Resumen

Problemas de especificación en SEM no lineal: la moderación que no lo era. **Antecedentes**: el análisis de efectos de interacción o moderación entre variables latentes es común en ciencias sociales. Sin embargo, cuando los predictores están correlacionados, los efectos de interacción y cuadráticos se vuelven parecidos, haciendo difícil su distinción. Así, cuando los datos provienen de un modelo de cuadrático a nivel poblacional y el modelo de análisis solo especifica interacciones, se pueden obtener resultados engañosos. **Método:** este artículo aborda las consecuencias de diferentes tipos de errores de especificación en modelos de ecuaciones estructurales no lineales utilizando un estudio de Monte Carlo. **Resultados:** los resultados muestran que estimar un modelo con interacciones cuando en la población hay efectos cuadráticos conducirá a una detección equivocada de efectos de moderación con casi plena seguridad, y lo mismo ocurrirá en el escenario opuesto. La estimación simultánea de interacciones y efectos cuadráticos en el modelo conduce a resultados correctos. **Conclusiones:** la estimación simultánea de efectos de interacción y cuadráticos permite evitar detectar efectos no lineales espurios o engañosos. Los resultados se discuten para ofrecer recomendaciones a los investigadores aplicados.

Palabras clave: ecuaciones estructurales no lineales, moderación, efectos de interacción, efectos cuadráticos, especificación del modelo.

Subject-matter theory often advises that relationships between variables may not be linear. For instance, inverted U-shaped relationships have been found between personality traits and job performance (Le et al., 2011), between positive affect and proactive behaviours (Lam, Spreitzer, & Frtiz, 2014), and between a wide range of psychological phenomena to do with well-being and performance (for an overview, see: Grant & Schwartz, 2011). These examples show that the relationship between two variables may form a parabola instead of a straight line, and such relationships are typically modelled as quadratic effects.

Even though quadratic effects are important both from a theoretical and practical point of view, interaction or moderation effects are the type of nonlinearity most commonly tested in applied research. For instance, interactions have been used to test the moderating effects of friendship on the relationship between victimisation and emotional maladjustment in victims of bullying (Barcaccia et al., 2018), and the moderating role of empathy in the relationship between guilt and antisocial behaviour (Barón, Bilbao, Urquijo, López, & Jimeno, 2018).

Interactions and quadratic effects are closely related concepts. While interactions reveal that the effect of one predictor depends on a second predictor, also known as a moderator variable, quadratic effects indicate that the effect of a predictor depends on the values of the predictor itself. Thus, quadratic effects may also be understood as an interaction between a variable and itself.

The majority of variables that are of theoretical interest in social and behavioural sciences are latent, and in recent years, nonlinear structural equation models (SEM) have grown in popularity among researchers seeking to estimate interaction or moderation effects between latent variables. However, given the similarities between interactions and quadratic effects, some authors have warned that during analysis of interactions, attention should also be paid to quadratic effects (Lubinski & Humphreys, 1990; MacCallum & Mar, 1995; Marsh, Wen, & Hau, 2006), as the two may become confounded if the predictors are correlated, leading to incorrect conclusions about the relationship between variables.

Received: August 7, 2019 • Accepted: September 27, 2019
Corresponding author: Karina Rdz-Navarro
Facultad de Ciencias Sociales
Universidad de Chile
775000 Ñuñoa, Santiago (Chile)
e-mail: rdznavarro@uchile.cl

One proposed—albeit controversial—solution is to fit models that estimate interaction and quadratic effects simultaneously. Some authors (Aiken & West, 1991; Shepperd, 1991) propose that nonlinear effects of any kind should only be included in the model on strong theoretical grounds. Others (Ganzach, 1997; Lubinski & Humphreys, 1990; Klein, Schermelleh-Engel, Moosbrugger, & Kelava, 2009) recommend simultaneous estimation of interaction and quadratic effects to avoid detection of spurious, misleading or overestimated interactions caused by incorrect model specification.

Harring, Weiss, and Li (2015) showed that simultaneous estimation of interaction and quadratic effects in latent variable models may indeed reduce the probability of detecting spurious interactions; however, this generates a major cost in terms of power for which the reduction in Type I errors would not compensate. The study stresses the consequences in terms of power of over-specified nonlinear models for populations in which either a single nonzero interaction effect or no nonlinear effect exists. It is unclear whether these results could be generalised to situations where, for example, quadratic effects exist within the population, but the analysis model only includes interaction effects.

Careful examination of this is particularly important, as our review of preliminary research revealed that more than 90% of published studies involving nonlinear SEM models only tested for two-way interaction effects and gave no consideration to quadratic effects. This shows that despite warnings concerning possible spurious or misleading interaction effects, applied researchers apparently tend to favour testing for a single type of nonlinear effect. In light of the debate surrounding nonlinear SEM model specification, further examination of the consequences of this practice is required.

The present study seeks to assess the effects of incorrect specification of nonlinear SEM models on parameter and standard error bias, Type I error and power in situations where interaction and/or quadratic effects exist in the population and the analysis model specification is correct (i.e., analysis model matches population model), misspecified (i.e., different type of nonlinear effect specified to that present), under-specified (i.e., fewer nonlinear effects specified than present) or over-specified (i.e., more nonlinear effects specified than present).

Overview of nonlinear SEM

Nonlinear SEM models were developed to estimate models with interaction effects (MI), quadratic effects (MQ), and interaction and quadratic effects simultaneously (MIQ), such as those presented in Equations (1), (2) and (3), where ξ_1 and ξ_2 are latent predictors; η is an endogenous factor; α is a latent intercept; ω_{12} is the slope of the interaction between ξ_1 and ξ_2 on η; ω_{11} and ω_{22} are the slopes of the quadratic effects of ξ_1 and ξ_2 on η, respectively; and ζ is the prediction error of η.

$$(MI)\ \eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \zeta, \tag{1}$$
$$(MQ)\ \eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{11}\xi_1^2 + \omega_{22}\xi_2^2 + \zeta, \tag{2}$$
$$(MIQ)\ \eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \omega_{11}\xi_1^2 + \omega_{22}\xi_2^2 + \zeta. \tag{3}$$

As with most SEM models, nonlinear SEM procedures assume that indicators reflect underlying factors, as shown in Equations (4) and (5), where x_i is the i^{th} exogenous indicator, y_i is the i^{th} endogenous indicator, τ_i represents the intercept of indicator i, λ represents the factor loadings of the indicators in their respective factor, and δ_i and ε_i represent random measurement errors.

$$x_i = \tau_i + \lambda_{ij}\xi_j + \delta_i, \tag{4}$$
$$y_i = \tau_i + \lambda_i\eta + \varepsilon. \tag{5}$$

Numerous nonlinear SEM procedures haven been proposed (e.g., Coenders, Batista-Foguet, & Saris, 2008; Kelava, Nagengast, & Brandt, 2014; Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Marsh et al., 2006), although the most popular nowadays is the latent moderated structural equations (LMS) method (Klein & Moosbrugger, 2000) which is readily implemented in Mplus (Muthén & Muthén, 2015). In contrast to traditional linear SEM methods based on analysis of limited information (e.g., variance-covariance matrices), the LMS method uses all of the information from a set of subject responses for model parameter estimation, and assumes that latent predictors and model errors (i.e., ξ_j, δ_i, ε_i and ζ) belong to normal distributions. This allows LMS to presume that deviations from a normal distribution of the endogenous factor η are the result of the nonlinear effects of the model. Thus, LMS will estimate the magnitude of the nonlinear parameters ω_{ij} using the Cholesky decomposition and the expectation-maximisation algorithm to produce maximum-likelihood estimates of model parameters (Klein & Moosbrugger, 2000).

Research has shown that, when distributional assumptions hold, the LMS method allows for unbiased and efficient parameter and standard error estimates, as well as Type I errors close to the nominal level (Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Kelava et al., 2014; Kelava et al., 2011). However, the literature also warns that, regardless of the method used for model estimation, the accuracy of results of nonlinear SEM models will be affected by two related factors: the degree of collinearity between predictors and the specification of the analysis model.

Collinearity of predictors and model specification

Research evidence has revealed that nonlinear SEM models are strongly affected by collinearity between predictors (Kelava, Moosbrugger, Dimitruk, & Schermelleh-Engel, 2008; Klein et al., 2009; MacCallum & Mar, 1995). In the simplest scenario, where two latent predictors ξ_1 and ξ_2 are mean-centred and normally distributed, the variances and covariances of the higher order terms ($\xi_1\xi_2$, ξ_1^2 and ξ_2^2) are defined only in terms of the variances and covariances of the original factors (Aiken & West, 1991); for example, $Cov(\xi_1\xi_2,\xi_1^2) = 2 \times Var(\xi_1) \times Cov(\xi_1,\xi_2)$ and $Cov(\xi_1^2,\xi_2^2) = 2 \times Cov(\xi_1,\xi_2)^2$.

Thus, when covariance between predictors is equal to zero, covariance between the nonlinear terms will also be zero, making them perfectly distinguishable from one another. However, as covariance between predictors increases, so does the relation between the products of latent variables, making them harder to differentiate. This higher collinearity could cause a variety of issues such as increases in the magnitude of standard errors—which reduces power (Harring et al., 2015)—or the detection of spurious nonlinear effects which would confound interactions with quadratic effects (and vice versa) when the analysis model is misspecified (Ganzach, 1997; Lubinski & Humphreys, 1990; Klein et al., 2009).

Indeed, parametric nonlinear SEM procedures require prior knowledge of the true structure of relations between factors in order to specify the model. This facilitates the testing of theory-driven hypotheses against the data, but also poses a challenge: how to anticipate the type of nonlinear effects that exist when specifying the model. Most social sciences studies that analyse nonlinear effects tend to estimate two-way interactions only (e.g., Bredow, 2015; Breevaart & Bakker, 2018; Masland & Lease, 2016). However, attention must also be paid to quadratic effects, especially if the factors are correlated (Klein et al., 2009; Marsh et al., 2006), owing to problems resulting from collinearity.

Some authors advise that nonlinear effects of any kind should only be included in the model if absolutely justified (Aiken & West, 1991; Shepperd, 1991), because each effect requires an additional significance test, resulting in increased Type I error rates. Researchers are therefore recommended to favour parsimonious models. Other authors (Cortina, 1993; Ganzach, 1997; Lubinski & Humphreys, 1990; Klein et al., 2009) claim that quadratic effects of the possible interacting variables should also be estimated, as interaction and quadratic effects share an important proportion of variance when predictors are correlated, and tend to confound each other. This means that, for example, if an interaction effect alone is estimated for a population in which quadratic effects exist, the probability of mistaking the latter for interactions increases as the correlation between predictors increases. The same is true in the opposite scenario. This points to estimation of potentially over-specified models as a means of compensating for the methodological problems caused by multicollinearity, and to the need to statistically check for possible confounders of the model's multiplicative effects in order to make robust inferences.

Although the need for control of Type I error rates is a concern common to these proposals, the advice they offer to applied researchers differs regarding how best to estimate nonlinear relations between variables. This is a particularly delicate issue given that goodness-of-fit statistics for nonlinear models are only recently being developed (e.g., Gerhard, Büchner, Klein, & Schermelleh-Engel, 2017; Gerhard et al., 2015), meaning that applied researchers must base their data analysis decisions on one of these approaches.

The debate surrounding estimation of potentially over-specified models began in the context of regression models for manifest variables. Most studies (e.g., Cortina, 1993; Ganzach, 1998) have used situations where either zero or nonzero interaction effects are present in the population to examine Type I error and power. Studies involving populations where non-zero quadratic effects are present and evaluating the consequences of model misspecification are less common. To the best of our knowledge, articles that have evaluated this phenomenon used regression models applied to variables with no measurement error (Ganzach, 1997; Ganzach, 1998; MacCallum & Mar, 1995), and showed that under-specified nonlinear models overestimate or incorrectly estimate interaction parameters, while misspecified models tend to detect spurious or misleading nonlinear effects, particularly as correlation between variables increases. However, given that these studies were conducted in the context of linear regression models, we must establish whether or not their results can be generalised to nonlinear SEM models.

Recent studies have extended the debate to nonlinear latent variable SEM models. Some suggest that the use of models which estimate interaction and quadratic effects together reduces the risk of inappropriate model selection (Gerhard et al., 2015; Klein et al., 2009). However, Harring et al. (2015) show that although over-specification reduces the probability of detecting spurious interaction effects, reductions in Type I error rates would be insufficient to compensate for greatly reduced power in detection of interaction effects.

Given the above—and in light of the widespread practice in applied research of testing interaction effects against data without consideration of quadratic terms—the present work contributes to determining the consequences of model specification errors under correlated predictors in nonlinear SEM models, and offers guidelines to researchers for correct estimation of nonlinear latent variable models.

Method

Procedure

A Monte Carlo study was conducted to create data for four types of structural model: a purely linear structural model (MLIN) based on Equation (6), a model with interaction effects (MI) based on Equation (1), a model using quadratic effects (MQ) based on Equation (2), and a model with interaction and quadratic effects (MIQ) based on Equation (3).

$$\text{(MLIN)}\ \eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \zeta \qquad (6)$$

In each condition, the exogenous factors ξ_1 and ξ_2 were simulated using a standard normal distribution. Prediction errors (ζ) were simulated using a normal distribution with mean zero. The variance of η was set to one and its mean was set to zero by adjusting the value of the latent intercepts (α). The γ parameters representing linear effects were fixed at .3 in order to represent a 9% explained variance of η.

The correlation between exogenous factors was set to .3 or .6 to represent moderately correlated and highly correlated factors, respectively. Given that the variance of the interaction depends on the degree of correlation between factors (Marsh et al., 2006), the interaction parameter for the MI and MIQ models was fixed at .219 and .192 when correlation was equal to .3 and .6, respectively. In both cases, the value assigned to the parameter equalled a 5% explained variance of η. The ω_{11} and ω_{22} parameters—representing quadratic effects—were fixed at .159 for the MQ and MIQ models, each representing a 5% explained variance of η.

The measurement models for the exogenous factors and for the endogenous factor were created based on Equations (4) and (5), respectively. Four indicators were simulated for each factor. The factor loadings of indicators of each factor (i.e., x_1-x_4, x_5-x_8, and y_1-y_4) were fixed at .8, .7, .6 and .5 to represent factors measured with a composite reliability of .75, and at .9, .85, .8 and .75 to represent factors with a composite reliability of .9. Sample sizes were set to 500 and 1,000 cases, representing medium and large samples. Thus, the present study worked with a total of 32 conditions in a 4 (population structural models) × 2 (correlation between predictors) × 2 (measurement reliability) × 2 (sample sizes) design. We generated 500 replicates for each condition.

Data analysis

The data produced by the four structural models were analysed using the MI, MQ and MIQ models shown in Table 1. The analyses using misspecified models were conducted to establish the Type I error rates resulting from model misspecification. The analyses using under-specified models were conducted to identify possible bias in parameter estimation resulting from omitted nonlinear effects. The analyses using correctly specified models were used to determine the power of the models. Finally, the analyses using over-estimated models were conducted to identify possible increases in Type I error rates for the linear structural model as a result of multicollinearity and/or an increase in the number of contrasts, as well as to establish the Type I error rates and the potential decrease in power for structural models with at least one type of nonlinear effect.

Each dataset was analysed using the LMS method in Mplus 7.4 (Muthén & Muthén, 2015). For each condition we analysed the proportion of convergent replicates and the bias in parameter estimates and standard errors. In conditions with population parameters not equal to zero we evaluated relative bias of parameter estimates, and in conditions with parameters equal to zero we evaluated the absolute bias (difference between the mean of estimates and zero). Relative and/or absolute bias in parameter estimates equal to or greater than |0.05| and relative bias in standard error estimates equal to or greater than |0.1| were considered unacceptable (Hoogland & Boomsma, 1998). We examined the proportion of replicates with significant results to a 95% confidence level in order to analyse Type I error rates and power for each analysis model. Following Bradley's liberal criterion (Serlin, 2000), Type I error rates of between .025 and .075 were considered acceptable. Power equal to or greater than .8 was considered acceptable (Muthén & Muthén, 2002).

Results

Optimal convergence rates were observed for all conditions, as well as biases below |0.05| for linear structural parameter estimates, and below |0.1| for standard error estimates. Factor correlations were estimated with irrelevant levels of bias in all conditions with the exception of the under-specified models. Parameters showing irrelevant levels of bias have been omitted from the results tables on the following pages.

As shown in Table 2, when the population involves only linear effects, use of MI, MQ and MIQ analysis models results in unbiased nonlinear parameter estimates and acceptable Type I errors. Using an over-specified analysis model when nonlinear effects are equal to zero does not seem to be advantageous in terms of control of Type I error rates in the detection of either interaction or quadratic effects. The results therefore refute the hypothesis that an increase in the number of contrasts and/or the multicollinearity of the model would lead to increases in Type I error rates.

Table 1
Population and data analysis models

Population model	Analysis model		
	MI	**MQ**	**MIQ**
MLIN	Over-specified	Over-specified	Over-specified
MI	Correctly specified	Misspecified	Over-specified
MQ	Misspecified	Correctly specified	Over-specified
MIQ	Under-specified	Under-specified	Correctly specified

Note: MLIN = purely linear model. MI = model with interaction effects. MQ = model with quadratic effects. MIQ = model with interaction and quadratic effects

When data were generated from a model with one interaction and the MI analysis model is used, unbiased results and optimal levels of power were achieved. When the analysis model was misspecified as a quadratic model, two misleading quadratic effects were obtained and detected as significant in 33%-93% of cases. The 5% variance represented by the interaction was detected as two small quadratic effects, each explaining a variance of between 1% and 1.5% of η when the population correlation between factors was .3 and .6, respectively. This implies that the true interaction effect is split into the estimated quadratic effects with a slight loss of information or explained variance. Incorrect detection of quadratic effects resulting from model misspecification increases with correlation between predictors or factor reliability. When the MI population model was analysed using an over-specified model, unbiased results were obtained and Type I error rates remained acceptable, except in one condition. However, statistical power varied substantially according to correlation between factors. When correlation was moderate, power to detect interaction effects remained at acceptable levels. When population correlation was high, power to detect interaction effects was reduced, reaching acceptable levels only when factor reliability was .9 and sample size was 1,000.

In the case of the quadratic population model, the correctly specified analysis model produced unbiased estimates with optimal levels of power. When the quadratic effects were misspecified as interactions in the analysis models, 90%-100% of replicates yielded significant results for interaction effects. Those effects represented an explained variance of between 4% and 5% of η when correlation between factors was moderate, and between 11% and 12% when correlation was high. This suggests that when quadratic effects are misspecified as interaction effects, not only will incorrect interactions be detected, but there will also be less information loss in the model's original proportion of explained variance. This may lead researchers to consider these models as more 'conclusive' due to the greater size of the detected effect. When the quadratic population model was analysed using an over-specified model, the parameters were estimated with no relevant bias. Interaction effects showed acceptable Type I error rates. Power to detect quadratic effects varied according to levels of factor reliability, correlation between factors, and sample size. Thus, when correlation between predictors is moderate, power is acceptable with sample sizes of 500 and factor reliabilities of .75. However, when correlation between predictors is high, acceptable power is possible at sample sizes of 500 and factor reliabilities of .9, or sample sizes of 1,000 and factor reliabilities of .75.

In population models with both interaction and quadratic effects, use of under-specified models produced the most critical over-estimation bias of nonlinear effects. Under-specification of models resulted in (a) over-estimation (by 6%-12%) of the correlation between factors only when this correlation was moderate, which was the case when MI and MQ (but not MIQ) models were estimated; and (b) over-estimation of nonlinear parameters in MI and MQ analysis models. In the MIQ population models simulated, total explained variance of nonlinear effects reached 15% (i.e., 5% explained variance for each nonlinear effect). However, over-estimation bias resulting from model under-specification implied that explained variance of estimated nonlinear parameters reached 18%-20% when correlation between factors was moderate, and 24%-33% when correlation was high. The results therefore imply that when nonlinear effects are omitted from analysis models, both

Table 2
Parameter bias and proportion of significant replicates

Analysis model			MI		MQ				MIQ					
			Bias	**Sig**	**Bias**		**Sig**		**Bias**			**Sig**		
ρ	n	Rel	ω_{12}	ω_{12}	ω_{11}	ω_{22}	ω_{11}	ω_{22}	ω_{12}	ω_{11}	ω_{22}	ω_{12}	ω_{11}	ω_{22}
Linear population model (all nonlinear effects are zero in the population)														
.3	500	.75	0.000[a]	.060	0.004[a]	-0.002[a]	.052	.040	0.000[a]	0.004[a]	-0.002[a]	.048	.050	.034
		.90	0.002[a]	.050	0.004[a]	0.001[a]	.068	.062	-0.001[a]	0.004[a]	0.001[a]	.052	.048	.046
	1,000	.75	-0.002[a]	.040	0.002[a]	0.001[a]	.064	.044	-0.006[a]	0.003[a]	0.003[a]	.058	.068	.046
		.90	-0.001[a]	.056	0.000[a]	-0.002[a]	.058	.054	0.002[a]	-0.001[a]	-0.003[a]	.054	.050	.062
.6	500	.75	-0.002[a]	.052	-0.002[a]	-0.004[a]	.066	.046	0.003[a]	-0.003[a]	-0.002[a]	.052	.046	.056
		.90	0.000[a]	.068	0.000[a]	-0.002[a]	*.078*	.058	0.006[a]	-0.003[a]	-0.004[a]	.052	.048	.046
	1,000	.75	0.001[a]	.046	-0.001[a]	0.001[a]	.050	.060	0.003[a]	-0.002[a]	0.000[a]	.064	.054	.070
		.90	0.000[a]	.039	0.002[a]	-0.002[a]	.072	.050	0.002[a]	0.001[a]	-0.003[a]	.054	.072	.060
Interaction population model (quadratic effects are zero in the population)														
.3	500	.75	-0.005	.970	**0.070[a]**	**0.071[a]**	*.330*	*.358*	0.000	0.000[a]	0.000[a]	.850	.046	.060
		.90	-0.014	1	**0.066[a]**	**0.063[a]**	*.482*	*.470*	-0.014	0.002[a]	-0.001[a]	.978	.060	.060
	1,000	.75	-0.023	.998	**0.072[a]**	**0.071[a]**	*.624*	*.638*	-0.032	0.002[a]	0.002[a]	.970	.044	.072
		.90	-0.027	1	**0.065[a]**	**0.063[a]**	*.726*	*.702*	-0.037	0.002[a]	0.000[a]	1	.046	.074
.6	500	.75	0.000	.982	**0.091[a]**	**0.091[a]**	*.474*	*.484*	-0.036	0.004[a]	0.003[a]	.312	.070	.052
		.90	0.005	.998	**0.087[a]**	**0.089[a]**	*.682*	*.746*	0.010	-0.003[a]	0.001[a]	.588	.066	.056
	1,000	.75	-0.005	1	**0.088[a]**	**0.093[a]**	*.760*	*.790*	-0.021	0.000[a]	0.004[a]	.492	*.080*	.068
		.90	-0.016	1	**0.086[a]**	**0.087[a]**	*.934*	*.934*	-0.026	0.001[a]	0.001[a]	.824	.042	.064
Quadratic population model (interaction effects are zero in the population)														
.3	500	.75	**0.208[a]**	*.900*	-0.019	-0.013	.958	.950	0.002[a]	-0.019	-0.019	.050	.880	.882
		.90	**0.187[a]**	*.934*	-0.013	-0.006	.998	.998	-0.001[a]	-0.006	0.000	.058	.978	.984
	1,000	.75	**0.209[a]**	*.994*	-0.013	0.006	1	1	-0.002[a]	-0.006	0.013	.066	.994	.992
		.90	**0.189[a]**	*1*	-0.006	-0.006	1	1	-0.001[a]	-0.006	0.000	.070	1	1
.6	500	.75	**0.298[a]**	*1*	-0.013	0.000	.920	.924	0.001[a]	-0.013	-0.006	.058	.552	.528
		.90	**0.287[a]**	*1*	-0.019	0.000	.994	.994	-0.002[a]	-0.013	0.006	.054	.848	.860
	1,000	.75	**0.299[a]**	*1*	-0.019	0.013	1	1	0.003[a]	-0.025	0.000	.044	.820	.830
		.90	**0.289[a]**	*1*	-0.013	0.000	1	1	0.003[a]	-0.019	-0.013	.054	.980	.988
Full model (both interaction and quadratic effects are present in the population)														
.3	500	.75	**0.977**	1	**0.478**	**0.434**	1	1	-0.009	0.013	-0.006	.868	.908	.916
		.90	**0.854**	1	**0.403**	**0.377**	1	1	-0.009	-0.006	-0.019	.990	.994	.992
	1,000	.75	**0.963**	1	**0.440**	**0.440**	1	1	-0.023	-0.006	-0.006	.988	.996	.998
		.90	**0.858**	1	**0.428**	**0.409**	1	1	-0.032	0.000	0.006	1	1	1
.6	500	.75	**1.578**	1	**0.579**	**0.572**	1	1	0.036	-0.006	-0.019	.414	.592	.546
		.90	**1.500**	1	**0.553**	**0.535**	1	1	-0.026	0.019	0.006	.708	.922	.948
	1,000	.75	**1.573**	1	**0.579**	**0.566**	1	1	0.000	0.006	0.000	.578	.850	.864
		.90	**1.500**	1	**0.535**	**0.547**	1	1	0.000	-0.006	0.000	.940	.998	1

Note: MI = model with interaction effects. MQ = model with quadratic effects. MIQ = model with interaction and quadratic effects. Sig = proportion of significant replicates. ρ = population correlation between ξ_1 and ξ_2. *n* = sample size. Rel = composite reliability of each factor. [a] = relative bias is not defined in this situation because the true population parameter is equal to zero; the mean deviation of estimates with respect to zero is displayed instead. Bold = over-estimated parameter. Italics = unacceptable level of Type I error rates according to Bradley's liberal criterion. Underlined = power less than .8

the parameters obtained and the model's total explained variance will be over-estimated. This only occurs with under-specified models.

When the analysis model was correctly specified, none of the estimated parameters showed relevant bias, but power to detect interaction and quadratic effects was seriously affected when correlation between factors was high. Thus, in order to achieve adequate power in the three estimated nonlinear parameters when correlation between factors is moderate, reliabilities of .75 and sample sizes of at least 500 are required, whereas to achieve the same power when factors are highly correlated requires reliabilities of at least .9 and sample sizes of at least 1,000. This is similar to when MI population models are analysed using MIQ models.

Regardless of the population model, power (although not Type I error) decreases when analysis models incorporate interaction and quadratic effects simultaneously. This reduction in power is caused neither by incorrect parameter estimation nor by standard errors, as in all cases these were estimated with acceptable levels of bias. In order to explore possible causes of a reduction in power in these situations, we assessed the heterogeneity of estimations by analysing the standard deviations of parameter estimates resulting from the three types of analysis model.

As shown in Table 3, the standard deviation of each nonlinear parameter estimate is higher in MIQ analysis models compared to MI and MQ. The magnitudes of the standard deviations are higher for interaction effects and increase as the correlation between predictors increases. Given that power to detect significant results depends on the variability of estimations, this greater heterogeneity could explain why power is lower in such situations and why Type I error is unaffected in the MIQ analysis model.

Conclusions

The present study analysed the consequences of model specification errors on the results of nonlinear SEM models. Two main conclusions may be drawn. Firstly, in the presence of correlated exogenous factors, incorrect specification (i.e., misspecification or under-specification) of nonlinear SEM models may result in the detection of incorrect and over-estimated nonlinear effects, which could have serious consequences for decision-making in research. For example, specification of an MI structural model when only true quadratic effects exist in the population will—in the majority of cases—result in over-estimated interaction effects being detected as statistically significant. Alternatively, if both interaction and quadratic effects are present in the population, estimation of a model with a single type of nonlinear effect will result in over-estimation of parameters and a spurious increase in total explained variance of the dependent variable.

Secondly, simultaneous estimation of interaction and quadratic effects would contribute to resolving these problems. Indeed, contrary to previous literature (Aiken & West, 1991; Shepperd, 1991), results presented here reveal that estimation of models that include interaction and quadratic effects simultaneously will not increase Type I error rates, will improve the precision of parameter estimation and will substantially reduce the detection of spurious nonlinear effects. Therefore, in order to avoid problems resulting from collinearity and incorrect structural model specification, it is advised that both interaction and quadratic effects be included simultaneously, even if the purpose of the study is to test for the existence of a single nonlinear effect. This is in line with other proposals concerning linear regression (Cortina, 1993; Ganzach, 1997; Lubinski & Humphreys, 1990) and with studies involving nonlinear SEM models (Gerhard et al., 2015; Klein et al., 2009).

That said, strong theoretical justification is still required for the inclusion of nonlinear terms in the model. Theory plays a key role in defining the requirement to estimate nonlinear models, but as the present study has demonstrated, researchers also need to consider methodological counterfactuals when specifying their models. Most applied studies analyse interaction effects but not quadratic effects, and the results of the present work suggest that a considerable number of these studies may have detected spurious moderation effects as a result of unmodelled quadratic terms in situations where predictors are correlated. As our knowledge is naturally incomplete and population models are by definition always unknown in any study involving real data, safeguards must be put in place. We therefore recommend estimation of nonlinear models that incorporate interaction and quadratic effects simultaneously, even if the aim is to test a theory that suggests the existence of only one type of nonlinear effect.

Spurious or misleading results are produced by incorrectly specified nonlinear models as a result of collinearity. The results presented here may therefore be generalised to, for example, models where more than two correlated predictors are analysed. However, in such scenarios a larger number of nonlinear effects (along with linear effects) would need to be modelled in order to avoid misleading or spurious results. More complex models may pose an even greater challenge in terms of factor reliability and the sample size needed to detect nonlinear effects. Indeed, consistent with findings of previous studies (e.g., Grewal, Cote, & Baumgartner, 2004; Kelava et al., 2014; Rdz-Navarro & Alvarado, 2015), results presented here reveal that large samples or scales with high reliabilities are needed to detect nonlinear effects. Future research should explore whether accurate results and acceptable power could also be achieved by, for example, increased sample sizes when factor reliabilities are lower, or increased factor reliabilities in smaller samples in situations where sizes of nonlinear effects are larger or smaller than those used in the present study. Researchers willing to estimate nonlinear effects between latent variables should carefully select highly reliable measurement instruments and collect the largest sample possible in order to ensure sufficient power to detect even small nonlinear effects.

Although these recommendations may sound restrictive and the model requirements (in terms of sample size and factor reliability)

Table 3
Standard deviations of parameter estimates by analysis model

Analysis model			MI	MQ		MIQ		
ρ	n	Rel	ω_{12}	ω_{11}	ω_{22}	ω_{12}	ω_{11}	ω_{22}
.3	500	.75	0.062	0.047	0.047	0.072	0.050	0.050
		.90	0.047	0.034	0.034	0.053	0.036	0.036
	1,000	.75	0.043	0.032	0.033	0.052	0.035	0.036
		.90	0.034	0.025	0.025	0.039	0.026	0.026
.6	500	.75	0.050	0.051	0.051	0.137	0.081	0.081
		.90	0.038	0.036	0.034	0.090	0.054	0.053
	1,000	.75	0.034	0.034	0.034	0.096	0.056	0.056
		.90	0.026	0.024	0.024	0.062	0.036	0.037

Note: MI = model with interaction effects. MQ = model with quadratic effects. MIQ = model with interaction and quadratic effects. ρ = population correlation between ξ_1 and ξ_2. *n* = sample size. Rel. = composite reliability of each factor difficult to meet, it is important to be both conscious and wary of the fact that the usefulness of SEM models for estimation of nonlinear effects depends on appropriate model specification. Simultaneous estimation of interaction and quadratic effects is therefore a useful means to guarantee adequate statistical inference.

Acknowledgements

This research was funded by the Chilean National Commission for Scientific and Technological Research (FONDECYT Grant 11160256).

References

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

Barcaccia, B., Pallini, S., Baiocco, R., Salvati, M., Saliani, A. M., & Schneider, B. H. (2018). Forgiveness and friendship protect adolescent victims of bullying from emotional maladjustment. *Psicothema*, *30*(4), 427-433. doi:10.7334/psicothema2018.11

Barón, M. J. O., Bilbao, I. E., Urquijo, P. A., López, S. C., & Jimeno, A. P. (2018). Moral emotions associated with prosocial and antisocial behavior in school-aged children. *Psicothema*, *30*(1), 82-88. doi:10.7334/psicothema2016.143

Bredow, C. A. (2015). Chasing prince charming: Partnering consequences of holding unrealistic standards for a spouse. *Personal Relationships*, *22*(3), 476-501. doi:10.1111/pere.12091

Breevaart, K., & Bakker, A. B. (2018). Daily job demands and employee work engagement: The role of daily transformational leadership behavior. *Journal of Occupational Health Psychology*, *23*(3), 338-349. doi:10.1037/ocp0000082

Coenders, G., Batista-Foguet, J. M., & Saris, W. E. (2008). Simple, efficient and distribution-free approach to interaction effects in complex structural equation models. *Quality & Quantity*, *42*(3), 369-396. doi:10.1007/s11135-006-9050-6

Cortina, J. M. (1993). Interaction, nonlinearity, and multicollinearity: Implications for multiple regression. *Journal of Management*, *19*(4), 915-922. doi:10.1016/0149-2063(93)90035-L

Ganzach, Y. (1997). Misleading interaction and curvilinear terms. *Psychological Methods*, *2*(3), 235-247. doi:10.1037/1082-989X.2.3.235

Ganzach, Y. (1998). Nonlinearity, multicollinearity and the probability of Type II error in detecting interaction. *Journal of Management*, *24*(5), 615-622. doi:10.1016/S0149-2063(99)80076-2

Gerhard, C., Büchner, R. D., Klein, A. G., & Schermelleh-Engel, K. (2017). A fit index to assess model fit and detect omitted terms in nonlinear SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(3), 414-427. doi:10.1080/10705511.2016.1268923

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: A Multidisciplinary Journal*, *22*(2), 276-287. doi :10.1080/10705511.2014.935752

Grant, A. M., & Schwartz, B. (2011). Too much of a good thing: The challenge and opportunity of the inverted U. *Perspectives on Psychological Science*, *6*(1), 61-76. doi:10.1177/1745691610393523

Grewal, R., Cote, J. A., & Baumgartner, H. (2004). Multicollinearity and measurement error in structural equation models: Implications for theory testing. *Marketing Science*, *23*(4), 519-529. doi:10.1287/mksc.1040.0070

Harring, J. R., Weiss, B. A., & Li, M. (2015). Assessing spurious interaction effects in structural equation modeling: A cautionary note. *Educational and Psychological Measurement*, *75*(5), 721-738. doi:10.1177/0013164414565007

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structural modeling: An overview and a meta-analysis. *Sociological Methods & Research*, *26*(3), 329-367. doi:10.1177/0049124198026003003

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. doi:10.1080/10705511.2014.915379

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. doi:10.1027/1614-2241.4.2.51

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. doi:10.1080/10705511.2011.582408

Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*(4), 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*(4), 647-674. doi:10.1080/00273170701710205

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

Lam, C. F., Spreitzer, G., & Fritz, C. (2014). Too much of a good thing: Curvilinear effect of positive affect on proactive behaviors. *Journal of Organizational Behavior*, *35*(4), 530-546. doi:10.1002/job.1906

Le, H., Oh, I. S., Robbins, S. B., Ilies, R., Holland, E., & Westrick, P. (2011). Too much of a good thing: Curvilinear relationships between personality traits and job performance. *Journal of Applied Psychology*, *96*(1), 113-133. doi:10.1037/a0021016

Lubinski, D., & Humphreys, L. G. (1990). Assessing spurious "moderator effects": Illustrated substantively with the hypothesized ("synergistic") relation between spatial and mathematical ability. *Psychological Bulletin*, *107*(3), 385-393. doi:10.1037/0033-2909.107.3.385

MacCallum, R. C., & Mar, C. M. (1995). Distinguishing between moderator and quadratic effects in multiple regression. *Psychological Bulletin*, *118*(3), 405-421. doi:10.1037/0033-2909.118.3.405

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

Masland, L. C., & Lease, A. M. (2016). Characteristics of academically-influential children: Achievement, motivation and social status. *Social Psychology of Education*, *19*(1), 195-215. doi:10.1007/s11218-015-9314-x

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus user's guide*. 7th Edition. Los Angeles, CA: Author.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(4), 599-620. doi:10.1207/S15328007SEM0904_8

Rdz-Navarro, K., & Alvarado, J. M. (2015). Reexamining nonlinear structural equation modeling procedures: The effect of parallel and congeneric measures. *Multivariate Behavioral Research*, *50*(6), 645-661. doi:10.1080/00273171.2015.1071236

Serlin, R. C. (2000). Testing for robustness in Monte Carlo studies. *Psychological Methods*, *5*(2), 230-240. doi:10.1037/1082-989X.5.2.230

Shepperd, J. A. (1991). Cautions in assessing spurious moderator effects. *Psychological Bulletin*, *110*(2), 315-317. doi:10.1037/0033-2909.110.2.315