MODELING THE MANIFEST COMPOSITES IN

LATENT NONLINEAR EFFECT MODELS

A Dissertation

by

YU-YU HSIAO

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Oi-Man Kwok
Committee Members,	Victor L. Willson
	Myeongsun Yoon
	Wen Luo
	Mark H. C. Lai
Head of Department,	Shanna Hagan-Burke

August 2017

Major Subject: Educational Psychology

ABSTRACT

Estimating nonlinear effects, including interaction and quadratic effects, is a prevailing issue in social and behavioral science. Despite several advantages of using latent variable models in estimating nonlinear effects, conducting path models with observed composites is still a very common practice among applied researchers. However, it is well-known among methodologists that conducting path models without considering the measurement errors of the observed composites would lead to biased estimation. Hence, the aim of this dissertation is to bridge the gap between methodologists and applied researchers by reviewing two methods— reliability adjusted product indicator (RAPI) and latent moderate structural equations (LMS)—which can be applied for estimating nonlinear effects while accounting for the measurement errors of the composites.

The dissertation is composed of three manuscripts. In the first manuscript I reviewed the RAPI and LMS methods and compared their performance with the conventional path models in terms of the estimation accuracy of the interaction effects. The second manuscript focuses on choosing the most appropriate reliability estimates while conducting the RAPI and LMS methods. In the third manuscript I discuss issues regarding having both interaction and quadratic effects in the models, and the impact of multicollinearity of the exogenous variables on the estimation of both nonlinear effects.

Based on the simulation results, I found that while estimating nonlinear effects with observed composites, conduct latent variable models and apply both the RAPI and

LMS methods yielded more accurate interaction estimates than the conventional path analysis. Additionally, for items following congeneric assumption, applying the RAPI and LMS methods with the Revelle's omega total yielded more accurate results; if only the power of the test is of interest, applying Cronbach's alpha, omega, and GLB make less difference. However, caution should be made for applying the RAPI and LMS methods when the correlation between the latent exogenous variables are high (i.e., over .5), especially when both interaction effects and quadratic effects are of interest. This dissertation concludes with a summary of findings and the implications of these findings in applied research.

To my grandmother, Lan-Hsiang Hsiao-Liang (蕭梁蘭香), for teaching me that family and friends are much more important than any paper. I miss you.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Kwok, and my committee members, Dr. Willson, Dr. Yoon, Dr. Luo, and Dr. Lai, for their input on this research. I am especially indebted to Dr. Kwok and Dr. Lai for all their support and advice throughout my doctoral journey.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience. A special thank you to the RMS group for creating a harmonic and supportive environment.

I would also like to thank my mother and father for the many years of financial support, and my roommate, Wilson Li, for being a fun person around and a supportive friend.

Finally, I would like to thank my wife, Rose, for her love, patience, and support throughout these five years, and our lovely son, Merlin, for being the light of our life. I could not have done this without you two.

CONTRIBUTORS AND FUNDING SOURCES

This work was supervised by a dissertation committee consisting of Professor Oi-Man Kwok [advisor], Professor Victor L. Willson, Professor Myeongsun Yoon, and Professor Wen Luo of the Department of Educational Psychology at Texas A&M University, and Professor Mark H. C. Lai of the College of Educational, Criminal Justice, & Human Services at the University of Cincinnati.

All work for the dissertation was completed by the student, under the advisement of of Professor Oi-Man Kwok of the Department of Educational Psychology at Texas A&M University and Professor Mark H. C. Lai of the College of Educational, Criminal Justice, & Human Services at the University of Cincinnati.

Graduate study was supported by the Graduate Assistant Scholarship and CEHD Strategic Research Award from the College of Education at Texas A&M University and a scholarship from the Ministry of Education, Taiwan.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

Structural Equation Modeling (SEM) is a common technique to estimate the effects with unobserved variables. Under the SEM framework, researchers can apply latent variable models by creating measurement-free variables, namely, latent variables from the observed variables, and define the corresponding measurement errors. Using latent variable models can increase estimation accuracy of the linear effects or nonlinear effects in the models.

Testing nonlinear effects including interaction and quadratic effects has been very important in social and behavioral science. Hence, methodologists have developed several latent variable models (e.g., the product-indicator approach for latent interaction effect) for modeling those nonlinear effects. Additionally, those latent variable models have been shown outperform the conventional path analyses, which did not take measurement errors into account, in terms of estimation accuracy and power.

Despite the advantages for using latent variable models, applied researchers generally used conventional path analyses in their research, especially for testing nonlinear effects such as interaction effects. One of the reasons is the commonly use of manifest composites instead of latent variables when analyzing nonlinear effects. Unfortunately, most of the latent variable models cannot be directly applied with manifest composite variables due to model identification issue (i.e., the under-identified model with only one observed indicator loaded on a latent factor which requires constraints to both factor loading and residual variance). This, in turn, leads to the inability to separate the latent variables and measurement errors from the observed composites.

In this dissertation, I conduct three studies to tackle this issue. In my first study, I compared two methods—reliability adjusted product indicator (RAPI) and latent moderate structural equations (LMS)—which can be applied to estimate the latent interaction effects with manifest composites, and evaluate the performance of these two methods with the conventional path analyses. One key feature in both the RAPI and the LMS methods is the use of the scale reliability to adjust for the measurement error variance of the exogenous composites. Therefore, in the second study, four commonly used reliability estimates: Cornbach's alpha, omega total, Revelle's beta, and greatest lower bond (GLB) were compared in terms of obtaining accurate and precise interaction effects estimates while incorporating with the LMS and the RAPI methods. In the third study, I extend the scope of my first two studies to quadratic effects, and further discuss issues including multicollinearity of the exogenous variables in the estimation of nonlinear effects. With the results and recommendations from these three studies, I intend to provide more feasible and effective approaches on testing non-linear effects to applied researchers whom will consider using latent variable models rather than traditional path models with manifest composites for estimating nonlinear effects.

CHAPTER II

EVALUATION OF TWO METHODS FOR MODELING MEASUREMENT ERRORS WHEN TESTING INTERACTION EFFECTS WITH OBSERVED COMPOSITE SCORES[1]

Introduction

Testing interaction effects is an important and common practice in social and behavioral research, as researchers are interested in determining whether the relationship between two variables stays the same or changes depending on the level of a third variable (i.e., the moderator). In practice, both the predictor and the moderator are measured by either a single item (e.g., socio-economic status, age, or gender) or a scale containing multiple items. For the applications of testing interaction effects with multiple-item exogenous variables, methodologists have proposed several statistical methods within the structural equation modeling (SEM) framework to test this type of interaction effects. These statistical methods are capable of modeling the latent interaction effects while simultaneously taking into account any measurement errors in the items (Jöreskog & Yang, 1996; Kenny & Judd, 1984; Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Lin, Wen, Marsh, & Lin, 2010; Little, Bovaird, & Widaman, 2006; Marsh, Wen, & Hau, 2004; Moulder & Algina, 2002; Wall & Amemiya, 2001).

---

[1] "Evaluation of Two Methods for Modeling Measurement Errors When Testing Interaction Effects With Observed Composite Scores" by Yu-Yu Hsiao, Mark H. C. Lai, and Oi-Man Kwok, 2017. *Educational and Psychological Measurement*. Copyright©2017 (SAGE). Reprinted by permission of SAGE Publications. DOI: 10.1177/0013164416679877

Despite methodological advancements in recent years, however, applied researchers still generally use observed composites (e.g., the mean or sum from a multiple-item scale) for both the predictor and the moderator when testing interaction effects. For example, a review of the papers (N = 120) published in the *Journal of Applied Psychology* in 2015 identified 22 (18.3%) articles testing at least one interaction effect using observed composites.[1] Of these 22 papers, only two corrected for the measurement errors of the exogenous variables, but in neither study did the author consider measurement errors in the interaction terms (Eby, Butts, Hoffman, & Sauer, 2015; Mitchell, Vogal, & Folger, 2015). In the remaining 20 (90.9%) articles, all the manifest variables and the corresponding interaction effects were assumed to be measured accurately (i.e., without any measurement errors). These findings echo those of Cole and Preacher (2014), who reviewed 44 issues of seven American Psychological Association journals published in 2011, and found that more than one tenth of the studies conducted path analyses without correcting for measurement errors in the manifest variables. Thus, ignoring measurement errors of the manifest variables and the corresponding interaction effects in path analyses is still quite common. Yet, perfectly reliable manifest variables rarely exist in real data (Cohen, Cohen, West, & Aiken, 2003) and, as a result, path analyses with observed variables uncorrected for measurement errors could result in biased (either under- or overestimated) path coefficients (e.g., Aiken & West, 1991; Busemeyer & Jones, 1983; Cole & Preacher, 2014) and lead to reduced statistical power (e.g., Marsh, Wen, Nagengast, & Hau, 2012).

Given the potential problems raised by failing to properly address measurement errors when observed composites are used, in this study, two alternative methods were reviewed and evaluated: the latent moderated structural equations (LMS) method and the reliability-adjusted product indicator (RAPI) method, both of which can properly take into account measurement errors when testing interaction effects based on observed composite measures. The LMS method, developed by Klein and Moosbrugger (2000), originally focused on testing interaction effect with multiple-indicator exogenous variables. In the present study, we illustrated how to impose error variance constraints on the exogenous variables while using the LMS method to estimate interaction effects based on observed composite variables. With regard to the RAPI method, even though it can be traced back to the 1980s (Bohrnstedt & Marwell, 1978; Busemeyer & Jones, 1983), it has seldom been used in applied research.

To our knowledge, the performance of these two alternative approaches in terms of the estimation accuracy of interaction effects with observed composites has yet to be investigated. Therefore, in the present study, we compared the LMS and the RAPI methods with the commonly used path analysis approach, which assumes no measurement error for all the observed composites and the corresponding interaction effect, under conditions of varying sample sizes, reliability levels, and magnitudes of the interaction effects.

Reliability Adjustment for the Interaction Effect between Observed Composites

As mentioned, the most common way to estimate interaction effects with observed composite scores is by using the traditional path models, assuming that all variables in the model are measurement-error free. Thus, under the traditional path model (see Figure 1), both the predictor and the moderator are presented as observed variables and are assumed to be measurement-error free. On the contrary, the distribution analytic method (see Figure 2) and the reliability-adjusted product indicator (RAPI) method (see Figure 3) can take into account the measurement errors of the exogenous variables while estimating interaction effects. A key feature of these alternative approaches is the application of a reliability adjustment of each observed composite by constraining the corresponding error variance. Below we first discuss how to impose the error-variance constraint with the use of reliability. We then present examples of applying these reliability adjustments to both LMS and RAPI methods.

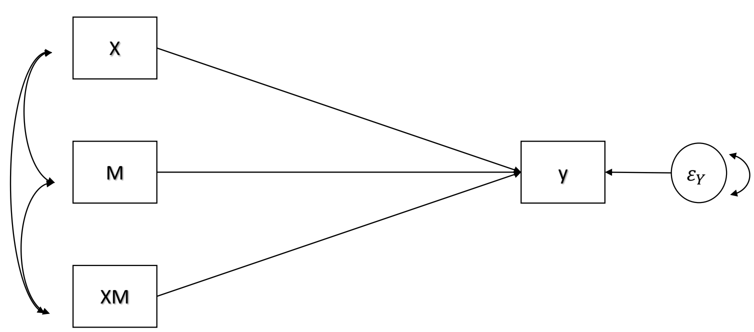

Figure 1. The path model for estimating one interaction effect with single predictor variable (X) and single moderator (M). Both X and M are composites from multiple items; XM is the product term of X and M.

In the classical testing theory (CTT) framework (Crocker & Algina, 1986; Lord & Novick, 1968), score reliability of a composite variable, *X*, is defined as the proportion of variance in *X* that can be attributed to the true score. Multiple approaches

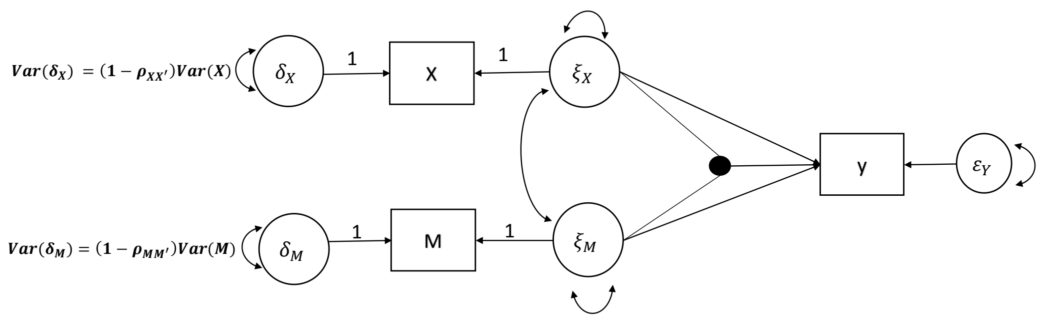

Figure 2. The latent moderated structural equations (LMS) method for estimating one interaction effect with single predictor variable (X) and single moderator (M). Both X and M are composites from multiple items. The equations for defining $var(\delta_X)$ and $Var(\delta_M)$ are cited from Bollen (1989).

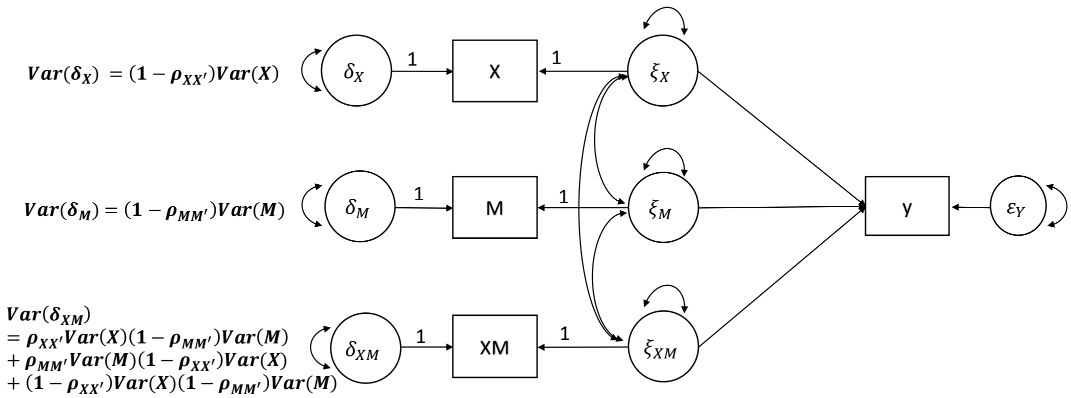

Figure 3. The reliability adjusted product indicator (RAPI) method for estimating one interaction effect with single predictor variable (X) and single moderator (M). Both X and M are composites from multiple items; XM is the product term of X and M. The equations for defining $Var(\delta_X)$ and $Var(\delta_M)$ are cited from Bollen (1989). The proof for defining $Var(\delta_{XM})$ is described in the Appendix A.

have been proposed to estimate reliability coefficients under conditions where the true-score variance cannot be directly obtained (Crocker & Algina, 1986). Among these approaches, structural equation modeling (SEM) is one of the techniques that yield more precise estimation of reliability coefficients (Raykov, 1997; Yang & Green, 2010). Let X_i be the *i*th observed item of a scale measuring the latent construct, ξ_X, with the measurement model written as below:

$$X_i = \tau_X + \lambda_{X_i}\xi_X + \delta_{X_i}, \quad (1)$$

where τ_X is the intercept, λ_{X_i} is the (unstandardized) loading of the *i*th indicator on ξ_X, and δ_{X_i} is the corresponding random measurement error term. Under the SEM framework, the factor structure reliability formula for this scale is written as (Bollen, 1989; Kline, 2011; Raykov, 1997; Raykov & Shrout, 2002):

$$\rho_{XX'} = \left(\sum \lambda_{X_i}\right)^2 Var(\xi_X) \Big/ \left[\left(\sum \lambda_{X_i}\right)^2 Var(\xi_X) + \sum Var(\delta_{X_i})\right], \quad (2)$$

where $Var(\xi_X)$ is the variance of the latent variable ξ_X and $Var(\delta_{X_i})$ represents the variance of the measurement error for the *i*th indicator.

If information about the individual item is unknown or unavailable (e.g., use of secondary data), one can only use the composite score, $X = \Sigma X_i$, as the single indicator for the latent variable, ξ_X. Thus, the corresponding reliability formula for X based on Equation (2) can then be rewritten as:

$$\rho_{XX'} = Var(\xi_X) \Big/ [Var(\xi_X) + Var(\delta_X)], \quad (3)$$

given that the only factor loading between X and ξ_X (i.e., λ_X) is constrained to 1.0 for identification purpose. Hence, the latent score ξ_X is equal to the true score in CTT (Borsboom, 2005). The error variance, $Var(\delta_X)$, can be estimated by using Equation (3), in which the reliability of a measure is the function of true-score variance and error variance as (Bollen, 1989):

$$Var(\delta_X) = (1 - \rho_{XX'})Var(X). \qquad (4)$$

Given the reliability coefficient, $\rho_{XX'}$, the error variance of X is a function of $(1 - \rho_{XX'})$, which is the proportion of the variance due to measurement error in X. The true score variance, $Var(\xi_x)$, can be rewritten as a function of the reliability coefficient and the observed variance, namely:

$$Var(\xi_X) = \rho_{XX'}Var(X). \qquad (5)$$

Equations (4) and (5) are the key elements in specifying the error variance constraints for the interaction effects under the RAPI method. Note that the discussion is equally applicable to mean composite scores, which is simply a rescaled version of the sum composite score.

Distribution analytic approach. Researchers can apply the distribution analytic approach to estimate interaction effects by either the latent moderated structural equations (LMS) method (Klein & Moosbrugger, 2000) or the quasi-maximum likelihood (QML) method (Klein & Muthén, 2007) under the SEM framework with specific data distributional assumptions. Figure 2 shows the simplest scenario in which a one-indicator predictor composite and a one-indicator moderator composite predict a single outcome. By using Equations (4) and (5) to constrain the error variances of the observed composites according to the corresponding reliability coefficient such as Cronbach's alpha (Bollen, 1989) or factor structure reliability (Raykov, 1997), one can estimate the latent interaction effect with the observed composite scores via the distribution analytic approach, which takes into account the measurement errors for the observed composites (Figure 2).

Based on Equations (4) and (5), $Var(\delta_X)$ and $Var(\delta_M)$ can, respectively, be defined as

$$Var(\delta_X) = (1 - \rho_{XX'})Var(X),$$

$$Var(\delta_M) = (1 - \rho_{MM'})Var(M),$$

while $Var(\xi_X)$ and $Var(\xi_M)$ can be defined as

$$Var(\xi_X) = \rho_{XX'}Var(X),$$

$$Var(\xi_M) = \rho_{MM'}Var(M) .$$

Although this is a very powerful approach, access to both the LMS and QML methods is quite limited. For example, the LMS method is exclusively built into Mplus

(Muthén & Muthén, 1998-2013) whereas the QML method is a stand-alone program available only from the developer Andreas Klein (Kwok, Im, Hughes, Wehrly, & West, 2016). Additionally, the overall model chi-square test and the commonly used model fit indices (e.g., CFI, RMSEA, and SRMR) are not available in these methods.

Reliability adjusted product indicator (RAPI) method. Researchers can also create a latent interaction effect factor by having the observed interaction effect term (i.e., the product of the predictor and the moderator) loaded on it (see Figure 3). Similar to the distributional analytic approach, the reliability-adjusted constraints can be directly applied to the exogenous variables (i.e., the predictor X and moderator M) under the RAPI approach, with the use of the same error-variance constraints as presented in Equations (4) and (5).

As for the observed interaction variable, XM, which is the product term of X and M, the variance of this interaction effect can be defined as the following equation (reproduced from Equation A7 in Appendix A), under the assumption of independent measurement errors and double mean-centered variables (Lin et al., 2010):

$$\begin{aligned} Var(XM) &= [E({\xi_{XM}}^2) - \left(E(\xi_{XM})\right)^2] + E({\xi_X}^2)E({\delta_M}^2) + E({\delta_X}^2)E({\xi_M}^2) \\ &\quad + E({\delta_X}^2)E({\delta_M}^2) \\ &= Var(\xi_{XM}) + Var(\xi_X)Var(\delta_M) + Var(\xi_M)Var(\delta_X) + Var(\delta_X)Var(\delta_M), \qquad (6) \end{aligned}$$

The procedure to create the double mean centered variable is straightforward. First both X and M are mean-centered, then the product term of the mean-centered X and M are mean-centered. The variance of the observed interaction variable, $Var(XM)$, can be decomposed into (a) the true- score variance, $Var(\xi_{XM})$, and (b) the error variance, $Var(\delta_{XM})$, which equals the last three components of Equation (6), or

$$Var(\delta_{XM}) = Var(\xi_X)Var(\delta_M) + Var(\xi_M)Var(\delta_X) + Var(\delta_X)Var(\delta_M). \quad (7)$$

The corresponding derivations are described in Appendix A. Accordingly, in Equation (6), we can substitute the measurement error variances and the true-score variances of X and M with their corresponding reliability estimates and observed variances. Hence, the error variance of the latent interaction effect is (Bohrnstedt & Marwell, 1978; Busemeyer & Jones, 1983) as follows:

$$Var(\delta_{XM}) = \rho_{XX'}Var(X)(1-\rho_{MM'})Var(M) + \rho_{MM'}Var(M)(1-\rho_{XX'})Var(X) + (1-\rho_{XX'})Var(X)(1-\rho_{MM'})Var(M). \quad (8)$$

Equation (8) is the key equation to set up the nonlinear constraint for the error variance of the latent interaction effect when using the RAPI method.

This study compared three methods of examining the interaction effects with observed composite scores to determine the estimation accuracy of the interaction effects. A Monte Carlo simulation study was conducted to compare methods with and without the consideration of measurement errors of the manifest variables. Both the LMS and RAPI methods were compared with the conventional path model. We chose the LMS method because it is currently the only distributional analytic approach that is feasible in a general SEM program (i.e., Mplus).

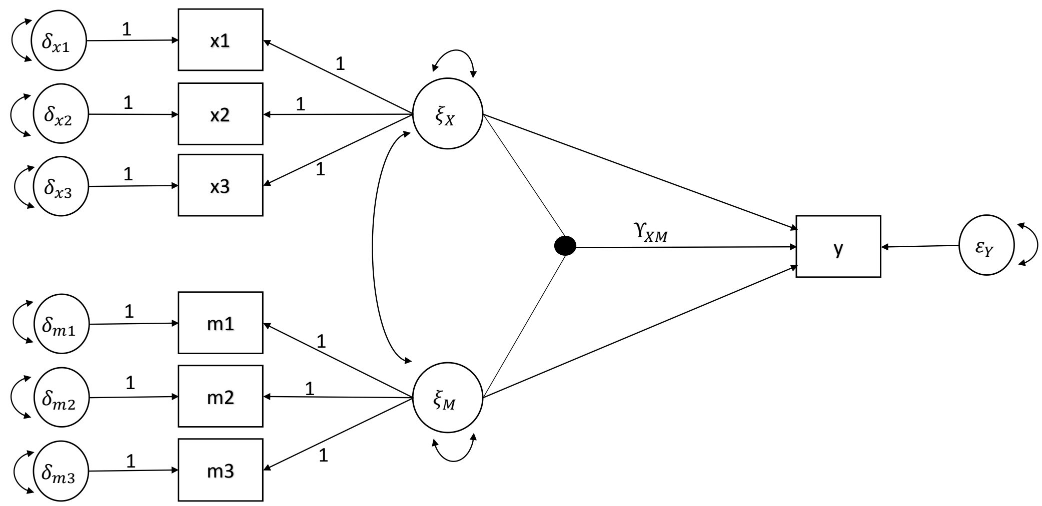

Figure 4. The pseudo population model with two latent exogenous variables and one observed variable.

Method

In this Monte Carlo study, we compared different methods for estimating the magnitude of the interaction effect Υ_{XM}, with the use of the data generation model shown in Figure 4. Specifically,

$$X_i = \tau_{X_i} + \lambda_{X_i}\xi_X + \delta_{X_i}, \quad (9a)$$

$$M_i = \tau_{M_i} + \lambda_{M_i}\xi_M + \delta_{M_i}, \quad (9b)$$

$$Y = \tau_Y + \Upsilon_X\xi_X + \Upsilon_M\xi_M + \Upsilon_{XM}\xi_{XM} + \epsilon_Y, \quad (9c)$$

where $X_i = X_1,\ X_2,\ X_3$ and $M_i = M_1,\ M_2,\ M_3$ were observed indicators, as shown in Figure 4. τ_{X_i}, τ_{M_i}, and τ_Y, respectively, represented the intercepts for X_i, M_i, and Y; all these intercepts were assumed to be zero. λ_{Xi} and λ_{Mi} were the factor loadings for the *i*th indicator on the two latent variables, ξ_X and ξ_M, respectively. δ_{X_i} and δ_{M_i} were the unique factors of the *i*th indicator on X_i and M_i, respectively. ξ_{XM} was the latent interaction variable between ξ_X and ξ_M. Finally, Υ_X, Υ_M, and Υ_{XM} were the path coefficients from the corresponding latent variables to the observed outcome Y, and ϵ_Y was the error term for Y. We chose a situation where mean composite scores were used in estimating the latent interaction effect. The results from this study are expected to be applicable to other forms of composite methods such as sum scores.

Monte Carlo Simulation Study

The model shown in Figure 4 was used to generate the population data. The two latent variables, ξ_X and ξ_M, and the two unique factors, δ_{x_i} and δ_{m_i}, were assumed to follow a standard normal distribution (i.e., mean equals to 0 and variance equals to 1.0) in the population. Both ξ_X and ξ_M were latent predictors with variance set at 1 and $Corr(\xi_X, \xi_M) = 0.5$. Υ_X and Υ_M were fixed to 0.3 (Evans, 1985). $Var(\epsilon_Y)$ was defined to make the variance of Y equal to 1 under the $\Upsilon_{XM} = 0$ condition. Therefore, $Var(\epsilon_Y) = 1 - (2 * 0.3^2 + 2 * 0.5 * 0.3^2) = 0.73$, indicating that the predictors as a whole explained 27% (large effect size; Cohen, 1988) of the variance in Y.

The items corresponding to ξ_X and ξ_M were assumed to be tau-equivalent items. Tau-equivalent items are defined as having equal loadings but possibly unequal error variance across items (Lord & Novick, 1968). Raykov (1997) showed that, if all the items (e.g., X_i and M_i in Figure 3 of the present study) under the common factor are tau-equivalent items, the estimated factor structure reliability equals Cronbach's alpha coefficient (Cronbach, 1951). In the present study, both λ_{Xi} and λ_{Mi} were fixed to 1.0. In terms of error variance of the exogenous variables, based on Equation (2), the sum of the error variances for the three items for each latent factor was 3.85 and 1.00, corresponding to .70 and .90 reliability, respectively. To achieve tau-equivalent items, we varied the error variances of the three items proportionally for both ξ_X and ξ_M. The error variance of the first item covered 55% of the total error variances in each latent predictor, followed by 33% of the second item, and 12% of the third item.[1] In other words, we manipulated the error variances as (2.12, 1.27, .46) for .70 reliability, and (.55, .33, .12) for .90 reliability. The design factors were described below.

Sample size, N. Based on the conditions used in past simulation studies (Cham, West, Ma, & Aiken, 2012; Chin, Marcolin, & Newsted, 2003; Lin et al., 2010; Marsh, Wen, & Hau, 2004; Maslowsky, Jager, & Hemken, 2015), we chose 100, 200, and 500 to represent small, medium, and relatively large sample sizes.

Reliability, ρ. We manipulated the reliability, ρ, for both X and M to be either .70 or .90. A Reliability of .70 represents 49% of the total variance being the true score variance and has been viewed as the acceptable lower boundary of reliability for group comparison in clinical research. Low reliability conditions (i.e., ρ < .70) were not considered in our simulation setting.

Interaction effect, Υ_{XM}. We manipulated the magnitude of the interaction effect Υ_{XM} to be either 0 (no interaction effect) or 0.50. The value of zero was designed to test the methods' performance when the null hypothesis was true (Cham et al., 2012). The value of .50 was used in a previous simulation study (cf. Chin et al., 2003).

Mplus 7.11 (Muthén & Muthén, 1998-2013) was used to generate 2,000 data sets for each condition. Given that the data were generated at the item level (i.e., three items per latent factor), we computed the mean composite scores for X and for M by averaging the corresponding items. Hence, we had three new observed composite scores; namely, the two observed composite variables X and M, and the corresponding product (or observed interaction effect) term XM. The data sets were then analyzed by fitting the three methods as shown in Figures 1, 2 and 3, respectively. For all three methods, double-centering strategy (Lin et al., 2010) was applied. Therefore, before analyzing the data using the three methods, X and M were first mean-centered; the product term XM was first computed using the mean-centered X and M and then mean- centered afterward. The annotated Mplus syntax for specifying the models with these three methods is presented in Appendix B.

Path model. The first method tested was the conventional path model (see Figure 1), with one predictor, one moderator, and the product term predicting one outcome variable. The measurement errors of the manifest exogenous variables were assumed to be zero. The three exogenous variables were allowed to be correlated.

LMS method. For the second method, the LMS method, no product indicator was created, as depicted in Figure 2. Instead, a maximum likelihood estimator with robust standard errors using numerical integration was used to estimate the latent interaction effect, based on the information of X and M. The measurement error variances for both X and M were constrained by using Equations (4) and (5). The two latent factors, ξ_X and ξ_M, were correlated. Both the common factor loadings were fixed to 1 for model identification purpose while the factor variances were freely estimated.

RAPI method. In the RAPI method, we utilized the reliability of each composite to constrain the corresponding measurement error. These non-linear constrains are shown in Figure 3. All the common factor loadings were fixed to 1 for model identification purposes whereas the factor variances were freely estimated. All the latent factors were allowed to be correlated.

Evaluation Criteria

Four criteria were applied to evaluate the performance of the three methods in examining the interaction effects with observed composite scores. The first two criteria, a 95% confidence interval (CI) coverage rate and the standardized bias, were used to evaluate bias – the average difference between the estimator and the true parameter. For the 95% CI coverage, the Wald interval was obtained, with a coverage rate > 91% considered acceptable (Muthén & Muthén, 2002). The standardized bias was the ratio of the average raw bias over parameter standard errors. Therefore, the standardized bias can be interpreted in a standard deviation unit, like Cohen's *d*. The standardized bias of the latent interaction effect estimates was compared with the cutoff value of 0.40. An absolute value < 0.40 was regarded as acceptable (Collins, Shafer, & Kam, 2001).

The third criterion was the relative standard error (SE) bias of the interaction effect estimates; it was designed to evaluate the precision of the interaction estimators. Estimators with smaller relative SE bias show less variability across simulation replications. As recommended by Hoogland and Boomsma (1998), relative SE bias values < 10% were considered acceptable.

Finally, the root mean square error (RMSE) was calculated to evaluate both the accuracy and precision of the parameter estimations for the three methods. The smaller the RMSE values, the more accurate the parameter estimations were across the 2,000 replications.

Results

The results of the conventional path model (without considering any measurement errors of the exogenous variables) and the models applying the RAPI and the LMS methods were compared in terms of the 95% CI coverage rate of the interaction effect, the standardized bias, relative standard error bias, and RMSE of the interaction effect estimates. The simulation results for $\Upsilon_{XM} = 0$ are displayed in Table 1 and the results for $\Upsilon_{XM} = 0.50$ are shown in Table 2.

Table 1

95% Confidence Interval (CI) Coverage Rate, Standardized Bias, Relative Standard Error (SE) Bias, and Root Mean Square Error (RMSE) for Υ_{XM} (=0) [a]

		95% CI Coverage (95%)			Standardized Bias			Relative SE Bias (%)			RMSE		
N	*ρ*	PM	RAPI	LMS	PM	RAPI	LMS	PM	RAPI	LMS	PM	RAPI	LMS
100	.70	93.7	97.0	91.5	-0.02	-0.01	-0.03	-3.69	-2.48	-9.7	0.07	0.19	0.12
	.90	94.0	94.1	91.0	-0.03	-0.03	-0.03	-4.57	-5.13	**-11.13**	0.08	0.10	0.10
200	.70	94.2	96.1	92.8	-0.03	-0.01	-0.03	-0.66	-2.56	-5.3	0.05	0.10	0.08
	.90	94.7	94.7	93.1	-0.03	-0.03	-0.04	-1.97	-2.21	-5.77	0.06	0.07	0.07
500	.70	94.6	94.1	93.8	-0.04	-0.03	-0.04	0.72	-0.91	-2.27	0.03	0.05	0.05
	.90	94.4	94.6	93.6	-0.04	-0.03	-0.03	-0.29	-0.49	-3.67	0.03	0.04	0.04

Note. *N* = sample size; *ρ* **=** reliability estimate**;** PM = Path model; RAPI = reliability-adjusted product-indicator method; LMS = latent moderated structural equations method.
[a]Values exceeding the recommended cutoffs are in boldface.

Table 2

95% Confidence Interval (CI) Coverage Rate, Standardized Bias, Relative Standard Error (SE) Bias, and Root Mean Square Error (RMSE) for Υ_{XM} (=0.5)[a]

		95% CI Coverage (95%)			Standardized Bias			Relative SE Bias (%)			RMSE		
N	ρ	PM	RAPI	LMS	PM	RAPI	LMS	PM	RAPI	LMS	PM	RAPI	LMS
100	.70	**13.4**	97.1	**90.0**	**-2.80**	0.30	-0.13	**-14.82**	1.14	-8.58	0.24	0.33	0.14
	.90	**79.2**	93.7	91.4	**-0.91**	0.07	-0.07	-9.63	-7.73	**-10.29**	0.12	0.11	0.11
200	.70	**0.0**	97.1	93.8	**-4.19**	0.26	-0.12	**-11.50**	0.28	-1.47	0.24	0.15	0.09
	.90	**67.9**	93.9	94.6	**-1.35**	0.05	-0.06	-6.48	-5.09	-2.88	0.10	0.07	0.07
500	.70	**0.0**	94.8	93.4	**-6.50**	0.16	-0.07	**-13.96**	-1.38	-1.60	0.23	0.08	0.06
	.90	**37.5**	93.3	93.8	**-2.11**	0.03	-0.04	-7.65	-6.26	-2.51	0.09	0.05	0.04

Note. *N* = sample size; ρ = reliability estimate; PM = Path model; RAPI = reliability-adjusted product-indicator method; LMS = latent moderated structural equations method.
[a]Values exceeding the recommended cutoffs are in boldface.

Convergence and Inadmissible Solutions

All the simulation replications were converged without any issues. Only 12 inadmissible solutions occurred with the RAPI method under the condition of non-zero interaction effect ($\Upsilon_{XM} = 0.50$), low reliability value ($\rho = .70$), and small sample size ($N = 100$). All 12 (out of 2,000 replications) non-positive definite matrices were due to the non-significant negative error variance in Y, accompanied with an inflated interaction effect Υ_{XM}. These 12 inadmissible solutions were excluded from the subsequent analyses. No inadmissible solution was found for either the conventional path model or the model using the LMS method.

Coverage of 95% CI of $\boldsymbol{\Upsilon_{XM}}$

As shown in Table 1, for conditions with interaction effect (Υ_{XM}) equal to zero, the coverage rate for the three methods were adequate, with a range from 93.7% to 94.7% for the conventional path model, from 94.1% to 97.0% for the RAPI method, and from 91.0% to 93.8% for the LMS method, regardless of sample size and the magnitude of reliability.

When the interaction effect was non-zero, the conventional path model without taking measurement errors into account generally resulted in lowest coverage rate. For example, as shown in Table 2, coverage rates were considerably low for the conventional path model, with a range from 0% to 79.2%. By comparison, under the same conditions, the coverage rates for the RAPI method continued to range from 93.3% to 97.1%. Similarly, the coverage rates for the LMS method were higher than those for the conventional path model, ranging from 90.0% to 94.6%. In other words, when the true interaction effect existed, the model that did not directly take measurement errors into account (i.e., the conventional path model) had the lowest chance of identifying the true effect.

Standardized Bias of Υ_{XM}

When the true interaction effect, Υ_{XM}, was set to zero, all three methods resulted in unbiased parameter estimates. That is, regardless of sample size and the magnitude of reliability, the standardized biases were adequate (i.e., |standardized bias| < 0.40): ranging from -0.04 to -0.02 for the path model, from -0.03 to -0.01 for the model utilizing the RAPI method, and from -0.04 to -0.03 for the model using the LMS method.

When the true interaction effect was not zero (= 0.50), the standardized biases of the interaction effects differed for the three methods across simulation conditions. For the conventional path model, substantial underestimations of the interaction effects were observed, with a range from -6.50 to -0.91 across all the conditions. By contrast, interaction effects were slightly overestimated for the RAPI method. These overestimations, however, were still within the acceptable criteria across all conditions. Standardized biases were larger (ranged from 0.16 to 0.30) under the low reliability (.70) condition, compared with those (ranged from 0.03 to 0.07) under the high reliability (.90) condition when using the RAPI method. On the other hand, slightly underestimated interaction effects were found for the LMS method, with standardized biases ranging from -0.13 to -0.07 under the low reliability (.70) condition, and from -0.07 to -0.04 under the high reliability (.90) condition.

Relative SE Bias of Υ_{XM}

As shown in Table 1, the absolute values of relative SE bias when $\Upsilon_{XM} = 0$ were all below 10% across all the simulation conditions for the conventional path model (ranged from -4.57% to 0.72%) and the model with the RAPI method (ranged from -5.13% to -0.49%). A negative SE bias indicates that the sample-estimated SE is, on average, smaller than the empirical standard error. Compared with the other two methods, the relative standard error biases were relatively higher for the LMS method. Additionally, under the high reliability (.90) and low sample size (100) conditions, the relative SE bias for the interaction effect estimates was the largest: -11.13% (i.e., underestimated by 11.13%). The relative SE biases for the other conditions from the LMS method ranged from -9.70% to -1.47%.

When $\Upsilon_{XM} = 0.50$, results of the relative SE biases varied among the three methods. As shown in Table 2, for the conventional path model, the relative SE biases were over 10% in absolute value (ranged from -14.82% to -11.50%) under the low reliability (.70) conditions regardless of sample size. The relative SE biases were below 10% in absolute value for all the conditions with high reliability (.90). For the RAPI method, all the relative SE biases were below 10% in absolute value. For the LMS method, the relative SE bias for the interaction effect estimates was -10.29% under the high reliability (.90) and small sample size (100) condition. For other conditions, the relative SE biases were all below 10% in absolute value (ranged from -8.58% to -1.47%). Although most of the SE biases for the RAPI and LMS methods were negligible, a trend of smaller SE bias in absolute value occurred for lower reliability (.70) conditions.

RMSE in Estimating Υ_{XM}

Generally, the RMSE values decreased as sample size or reliability increased. Under the condition of $\Upsilon_{XM} = 0$, the RMSE values were the highest with the RAPI method (ranged from 0.04 to 0.19), followed by the LMS (ranged from 0.04 to 0.12) method and the path model (ranged from 0.03 to 0.08).

On the other hand, different RMSE patterns were observed when $\Upsilon_{XM} = 0.5$, in which the RMSEs of the PM method were overall the highest across all three methods. One exception was when the sample size was small (100) and the reliability was low (.70), here the RMSE of the parameter estimates under the RAPI methods (RMSE = 0.33) was higher than that of the path model (RMSE = 0.24). For all the other simulation conditions, the RMSEs for both RAPI and LMS methods were lower than those from the path model. Overall, the parameter estimates yielded from the LMS method were the most precise and accurate (i.e., RMSE ranged from 0.04 to 0.14) among the three methods. Finally, sample size had less influence on the RMSE values of the path model.

Discussion

Despite the existence of the SEM approach for decades, applied researchers still commonly test interaction effects with the presumably measurement-error-free observed composite scores. In this study, we reviewed two alternative methods, namely, the reliability adjusted product indicator (RAPI) method and the latent moderated structural equations (LMS) method, and compared their performance with that of the conventional path model through a Monte Carlo study.

Our simulation results showed a substantial negative standardized bias and considerably low coverage rate when the conventional path model (without adequately taking into account measurement errors of the observed composites) was employed in testing interaction effect. Thus, the interaction effect under the conventional path model is more likely to be underestimated from the true population value when measurement errors are not adequately taken into account in the analysis. These findings reaffirm past research, which has shown biased results due to imperfect (reliability) measurement when testing interaction effects (Dunlap & Kemery, 1988; Evans, 1985; Feucht, 1989). Thus, the conventional path models, which do not adjust for measurement errors of the manifest predictors, are not recommended for testing interaction effects.

On the other hand, the two alternative methods discussed here, namely, the RAPI and LMS methods, can directly adjust the measurement errors of the observed composites by using either the factor structure reliability calculated from the measurement model or the conventional coefficient alpha. The major difference between these two methods is how the interaction effect is specified/captured: RAPI requires the creation of a product indicator for the latent interaction effect, whereas LMS does not. Results from the present study have shown that the RAPI method performed comparably well to the LMS method in estimating the interaction effects. Additionally, when the true interaction effects were non-zero, RAPI yielded slightly over-estimated (but still acceptable) coefficients, whereas LMS yielded slightly underestimated coefficients.

Hence, the LMS method may be more preferable for applied researchers who aim to be more conservative by preventing overestimated effects.

Both sample size and the magnitude of reliability played important roles in estimating the non-zero interaction effect. The standardized biases became smaller as sample size increased for both RAPI and LMS methods, suggesting that the reliability-adjusted measurement error constraints worked better with larger sample sizes. Reliability had a similar effect on standardized biases. With the same sample size, higher reliability (.90) produced more accurate interaction effect estimates than those from lower reliability (.70). Additionally, the RAPI method yielded less stable estimates than the LMS method under the low reliability and small sample size condition. Hence, the LMS method is more preferable when the exogenous variables are less reliable along with a small sample (e.g., $N = 100$).

Although our simulation results showed the benefits of controlling for measurement errors when testing interaction effects, this step sometimes comes at the price of increasing variability. For example, comparing four latent interaction modeling approaches, Cham and colleagues (2012) found that latent variable models can correct for bias but sometimes lose statistical power. When estimating the non-zero interaction effects in our simulation, the relative SE biases of the interaction effects from RAPI and LMS were higher than those from the path model under the high reliability (.90) condition. Given the reciprocal relationship between measurement error and reliability, these results suggest that constraining measurement errors for highly reliable variables may lead to over-correction, especially when the sample size is small. However, if we consider precision and bias together, the RMSE results showed that both the RAPI and LMS methods in general outperformed the conventional path model. Hence, these measurement-error adjustment methods are recommended for testing interaction effects with composites, with the recognition that the RAPI method may produce less precise or less accurate estimates than the LMS method under conditions with small sample and less reliable measures.

Practically speaking, there are several situations where researchers will find both the RAPI and LMS methods more preferable than the multiple-item latent factor model in empirical data analyses. For example, if the predictors or the moderators are measured by a large number of items, fitting the hypothesized structural model at the item level may lead to convergence issues due to the complexity of the model.

Another example would be when researchers analyze secondary data and have limited or no access to the original items. As mentioned earlier, the factor structure reliability in SEM is comparable to the conventional internal consistency reliability (i.e., Cronbach's alpha or coefficient alpha) with tau-equivalent items (i.e., items with equal factor loadings and possibly unequal error variances). Hence, as long as the reliability information of the composites is available, we advocate the use of this information to constrain the error variances for the observed composites and conducting the analyses with either the RAPI or LMS method to obtain interaction effect estimates.

Limitations and Future Research Directions

Two limitations in the present study must be addressed. First, since the interaction effect is the product term of the predictor and moderator, having a low reliability on either or both variables can amplify the measurement error of the interaction effect (Aiken & West, 1991). It is, therefore, worth investigating how changes in the reliability of the interaction term influence the interaction effect estimation. Second, the scope of this study was the traditional single-level interaction effect. Future study is needed to investigate the impact of ignoring measurement errors when testing interaction effect with observed composites under more complex data structures, such as multilevel data.

CHAPTER III

COMPARISONS OF RELIABILITY ESTIMATES FOR CORRECTING MEASUREMENT ERRORS OF THE EXOGENOUS COMPOSITES WHEN TESTING INTERACTION EFFECTS

Introduction

Social and behavioral research often relies on interaction effects, which indicate the direction and magnitude of the relation between exogenous variables and endogenous variables. As being shown in Chapter 2, the variables are often measured with less than perfect reliability and lead to biased estimates of the interaction effects. Latent variable models under the Structural Equation Modeling (SEM) framework can effectively mitigate the biased estimation by creating error-free latent variables to replace the non-perfectly measured observed variables. In decades, numerous methods under the SEM framework have been proposed to incorporate latent variable models in estimating interaction effects (Jöreskog & Yang, 1996; Kenny & Judd, 1984; Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Lin, Wen, Marsh, & Lin, 2010; Little, Bovaird, & Widaman, 2006; Marsh, Wen, & Hau, 2004; Moulder & Algina, 2002; Wall & Amemiya, 2001).

Many studies involve interaction effects in the social and behavioral sciences use composite scores from multiple items. Such practices result in challenging scenario for researchers who conduct latent variable models while estimating the measurement errors of the predictors. For example, having both latent factors and latent measurement error variance related to single indicator (i.e., the composite) being freely estimated is not legitimated in SEM. One solution is to use the reliability of the composites to make the “best guess” of the measurement error variance before running the model (Bollen, 1986, Aiken & West, 1996).

In the practices of estimating interaction effects with composite variables, in the previous chapter, I evaluated two methods under the latent variable models– latent moderated structural equation (LMS) and reliability-adjusted product indicator (RAPI) methods—which can be used for estimating the interaction effects of composite variables while accounting for measurement errors of the predictors. These two methods utilized the strategies proposed by Bollen (1986) and Bohrnstedt & Marwell (1978) to pre-set the measurement error variance of the exogenous composites, which were double-mean-centered (Lin et al., 2010). In the previous chapter I have demonstrated that both the LMS and RAPI methods yielded less biased estimation of the interaction effects, compared with the conventional path analyses.

Despite the promising results, two questions have arisen. First, the measurement structures in the population model were defined as tau-equivalence items, which assuming factor loadings are invariant whereas error variance are varied across items. Such assumption is hard to achieve in real-world data (Green & Yang, 2009) Therefore, whether these two methods perform equally well with the congeneric equivalent items (i.e., invariance factor loadings and error variance) has yet to be investigated.

Secondly, one important feature in both LMS and RAPI methods is to constrain the measurement error variance of the composite exogenous variables by using the reliability of the item scores (Aiken & West, 1991; Bohrnstedt & Marwell, 1978; Bollen, 1989). In chapter 2, the omega reliability (McDonald, 1978) was used and were assumed to be equivalent to Cronbach's alpha with tau-equivalent items. However, under the item assumption of congeneric equivalence, omega has found to be substantially different from Cronbach's alpha (Rayko, 1997). Additionally, previous researchers have proposed several alternative reliability values other than omega estimates for the Cronbach's alpha. Therefore, the present study investigates the differences in the estimation of the interaction effects with different reliability formula being used. Based on the popularity in literature and the ease to conduct with accessible programs, we compared four reliability estimates: Cronbach's alpha, omega total, Revelle's omega total, and greatest lower bound.

Interaction Effects with Composite Scores

When the effects of two latent variables (ξ_X and ξ_M) and their interaction effect (ξ_{XM}) on an endogenous variable η_Y (η_Y is a latent factor variable tapping multiple items) is considered, the following latent variable model has been used to estimate the interaction effects:

$$\eta_Y = \tau + \gamma_X \xi_X + \gamma_M \xi_M + \gamma_{XM} \xi_{XM} + \varepsilon_Y, \quad (1)$$

where τ is the intercept, γ_X and γ_M represent the linear effects, γ_{XM} represents the interaction effect, and ε_Y is the disturbance of η_Y. Assuming ξ_X, ξ_M and ξ_{XM} each measured by composite variables *X*, *M*, and the product term *XM*, respectively. The measurement models for ξ_X, ξ_M and ξ_{XM} can be described as:

$$X = \tau_X + \lambda_X \xi_X + \delta_X, \tag{2a}$$

$$M = \tau_M + \lambda_M \xi_M + \delta_M, \tag{2b}$$

$$XM = \tau_{XM} + \lambda_{XM} \xi_{XM} + \delta_{XM}, \tag{2c}$$

where λ_X, λ_M, and λ_{XM} are equal to 1 for model identification purpose. Likewise, additional constraints should be made on the variance of δ_X, δ_M, and δ_{XM}. One common strategy is to use the reliability and variance of *X* and M to preset the variance of δ_X, δ_M, and δ_{XM} (Bollen, 1989; Bohrnstedt & Marwell, 1978). Specifically,

$$Var(\delta_X) = (1 - \rho_{XX'})Var(X), \tag{3a}$$

$$Var(\delta_M) = (1 - \rho_{MM'})Var(M), \tag{3b}$$

$$Var(XM) =$$

$$Var(\xi_{XM}) + Var(\xi_X)Var(\delta_M) + Var(\xi_M)Var(\delta_X) + Var(\delta_X)Var(\delta_M) \tag{3c}$$

where $Var(.)$ represents the variance component. $\rho_{XX'}$ and $\rho_{MM'}$ are the reliability estimate of the item scores of *X* and *M*. Note that equation (3c) can only be hold when both *X* and M follow bi-normal distribution (Bohrnstedt and Marwell, 1978; Busemeyer and Jones, 1983) or are double mean-centered variables (see Appendix A).

Methods for Estimating the Reliability of Scales

In literature, there are over 30 methods which can be applied to estimate the reliability of scales (Hattie, 1985). In the present study, we focus on Cronbach's alpha, omega, Revelle's beta, and greatest lower bond (GLB) based on (1) the methods' popularity among substantive studies, (2) conceptually similar in terms of reliability, and (3) the accessibility of computer program/package.

Cronbach's alpha ($\alpha lpha$). Cronbach's alpha (Cronbach, 1951) is no doubt the most prevailing reliability formula among social and behavioral research nowadays. Specifically,

$$\alpha lpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} s_i^2}{s_X^2}\right), \tag{4}$$

where k is the number of items, s_i^2 is the variance of individual item *i* where *i* = 1, …, *k*, and s_X^2 is the variance of the items' total scores on the scale. Alpha will reach its maximum when the ratio of the sum of the variance of individual items over the variance of the items' total scores close to zero, indicating that respondents provide similar answers to a set of items which are designed under the same domain. In such situation, the set of items would be considered to have high reliability in terms of internal consistency (Crocker & Algina, 2008).

Several criticisms have arisen of alpha regarding measuring the lower bond of the reliability (e.g., Green & Yang, 2009; Revelle & Zinbarg, 2009; Sijtsma 2009). Alpha has been known to underestimate the reliability for several occasions. For example, Sijtsma (2009) compared Cronbach's alpha with several reliability estimates and conclude that the greatest lower bond (GLB) is the best reliability estimate. Revelle & Zinbarg (2009) argue that GLB is not the best by including the reliability estimates being compared in Sijtsma (2009) and omega total (McDonald, 1978). They found that omega total in general yielded higher reliability values than both alpha and GLB. Given the controversial findings among the alternatives of Cronbach's alpha, further investigation on the Omega total and GLB is needed.

Omega total (ω_{Total}). Omega total (McDonald, 1978) was calculated after conducting a confirmatory factor analysis (CFA) on a measurement model of a scale. Such calculation was based on the CFA parameters. For a model without error covariance, omega total (ω_{Total}) of a scale is estimated as follows:

$$\omega_{Total} = \frac{\left(\sum_{i=i}^{k} \lambda_i\right)^2 \phi}{\left(\sum_{i=i}^{k} \lambda_i\right)^2 \phi + \sum_{i=1}^{k} \theta_{ii}}, \quad (5)$$

where λ_i is the standardized or unstandardized factor loading for the ith item on the scale, ϕ is the estimated factor variance, θ_{ii} is the error variance for the ith item, and k is the number of items on the scale. Equation (5) can be extend to fulfill conditions with at least one error covariance exist (Raykov, 1997). Specifically,

$$\omega_{Total} = \frac{\left(\sum_{i=i}^{k} \lambda_i\right)^2 \phi}{\left(\sum_{i=i}^{k} \lambda_i\right)^2 \phi + \sum_{i=1}^{k} \theta_{ii} + 2\sum_{i=1}^{k} \theta_{ij}}, \quad (6)$$

In the present study, the error variance of each indicator was assumed to be independent. Hence, we focus on ω_{Total} calculated through Equation (5) in the present study.

Greatest Lower Bound (GLB). Jackson and Agunwamba's (1977) greatest lower bond (GLB) to reliability is another alternative to the Cronbach's alpha which being discussed among methodologists. Sijtsma (2009) explain the GLB as follows. The item observed covariance matrix C_X can be decomposed into the sum of the item true score matrix C_T and the error covariance matrix C_E, namely,

$$C_X = C_T + C_E, \tag{7}$$

where all three matrices are positive semi-definite (psd) which cannot have a negative eigenvalue. Since C_E and C_T are estimated and conditional on each other. Jackson and Agunwamba (1977) focus on creating all the possible set of C_E which allow $C_X - C_E$ has no negative eigenvalue. GLB is defined by utilizing all the solutions form Equation (7), specifically,

$$GLB = 1 - \frac{trace[C_E]}{S_X^2}, \tag{8}$$

where S_X^2 is the variance of the observed items and $trace[C_E]$ represent the maximal values for the possible measurement error matrix. Hence, GLB indicates the lowest possible value of reliability from the data (Bentler & Woodward, 1980). Likewise, when the glb for a scale is 0.8, the true reliability will be within the range of 0.8 and 1.

Revelle's Omega. As opposed to the model used for estimating omega total, the model being specified to calculate Revelle's omega is a bifactor model, in which each item is influenced by a general factor and group factor(s). Specifically,

$$\omega_{RT} = \frac{(\sum_{i=1}^{k} \lambda_{gi})^2 + (\sum_{f=1}^{F} \sum_{i=1}^{k_f} \lambda_{fi})^2}{V_X}, \tag{9}$$

where λ_{gi} is the loading of the *i*th item on the general factor, λ_{fi} is the standardized loading of the *i*th item on the *f*th group factor, *k* is the total number of items, *F* is the total number of group factors, and k_f is the number of items that load on the fth group factor. V_X is the total variance after rotation which is equal to the sum of each element of the sample correlation matrix. Conceptually, Equation (9) is equal to Equation (5) which defines reliability as the ratio between true score variance and total score variance. However, the variance components (including the factor loadings) are estimated by using the *Schimd-Leiman rotation* (Schmid & Leiman, 1957) and may lead to different reliability estimates from the omega total (Revelle, 2016).

Purpose of the Study

This study investigated the impact of using different reliability coefficients to adjust for the measurement error variance of the exogenous composites while estimating interaction effects. A Monte Carlo simulation study was conducted to compare four reliability estimates, including Cronbach's alpha, omega total, Revelle's omega total, and greatest lower bond. The accuracy, precision, and power of the interaction estimations over sample size, and the levels of reliability between the RAPI and the LMS methods would be investigated.

Method

In this Monte Carlo study, we compared the LMS and RAPI methods with four different reliability calculations for estimating the magnitude of the interaction effect Υ_{XM}, with the use of the data generation model shown in Figure 1. Specifically,

$$X_i = \tau_{X_i} + \lambda_{X_i}\xi_X + \delta_{X_i}, \tag{10a}$$

$$M_i = \tau_{M_i} + \lambda_{M_i}\xi_M + \delta_{M_i}, \tag{10b}$$

$$Y_i = \tau_Y + \lambda_{Y_i}\eta_Y + \epsilon_{Y_i}, \tag{10c}$$

$$\eta_Y = \tau_Y + \Upsilon_X\xi_X + \Upsilon_M\xi_M + \Upsilon_{XM}\xi_{XM} + \varepsilon_Y, \tag{10d}$$

where $X_i = X_1,\ X_2,\ X_3$ and $M_i = M_1,\ M_2,\ M_3$ were observed indicators, as shown in Figure 1. τ_{X_i}, τ_{M_i}, and τ_Y, respectively, represented the intercepts for X_i, M_i, and Y; all these intercepts were assumed to be zero. λ_{Xi}, λ_{Mi}, and λ_{Yi} were the factor loadings for the *i*th indicator on the three latent variables, ξ_X, ξ_M, and ξ_Y respectively. δ_{x_i} and δ_{m_i} were the unique factors of the *i*th indicator on X_i and M_i, respectively. ϵ_{Y_i} was the unique factor of the *i*th indicator on Y_i. ξ_{XM}, as shown as a black dot in Figure 1, was the latent interaction variable between ξ_X and ξ_M. Finally, Υ_X, Υ_M, and Υ_{XM} were the path coefficients from the corresponding latent variables to the observed outcome η_Y, and ε_Y was the error term for Y. We chose a situation where mean composite scores were used in estimating the latent interaction effect. The results from this study are expected to be applicable to other forms of composite methods such as sum scores.

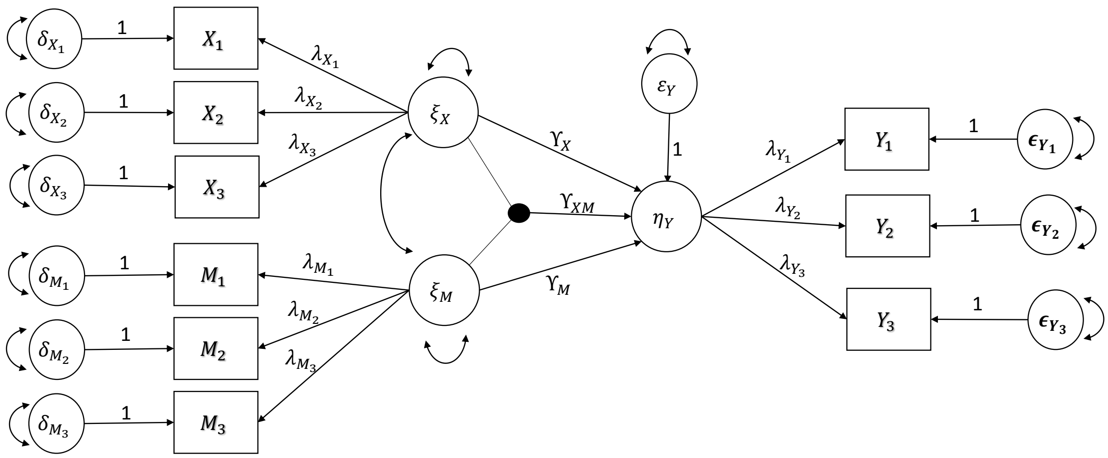

Figure 5 The pseudo population model with two latent exogenous variables and one latent endogenous variable. Each latent variable directly influences three indicators.

Monte Carlo Simulation Study

The model shown in Figure 5 was used to generate the population data. The latent true score variables ξ_X, ξ_M, and η_Y were assumed to follow a standard normal distribution (i.e., mean equals to 0 and variance equals to 1.0) in the population. $Corr(\xi_X, \xi_M) = 0.3$. The items corresponding to ξ_X and ξ_M were assumed to be congeneric items. Congeneric items are defined as having factor loadings and error variance vary across items (Jöreskog, 1971; Millsap & Everson, 1991). In the present study, the loadings for the first indicator of each latent true score (i.e., λ_{X_1}, λ_{M_1}, and λ_{Y_1}) were fixed to 1.0. The other two loadings were randomly selected from a uniform distribution range from 0.7 to 1.0. Hence, λ_{X_2}, λ_{M_2}, and λ_{Y_2} were fixed to 0.89 and λ_{X_3}, λ_{M_3}, and λ_{Y_3} were fixed to 0.72.

The linear effects Υ_X and Υ_M were fixed to 0.3 (Evans, 1985) and the interaction effect Υ_M was fixed to 0.2. $Var(\varepsilon_Y)$ was defined to make the variance of η_Y equal to 1 under the $\Upsilon_{XM} = 0$ condition. Therefore, $Var(\varepsilon_Y) = 1 - (2 * 0.3^2 + 2 * 0.3 * 0.3^2) = 0.766$, indicating that the predictors as a whole explained about 23% (large effect size; Cohen, 1988) of the variance in Y. The design factors were described below.

Sample size, $\boldsymbol{N}$. In study 1 I reviewed the conditions used in past simulation studies (Cham, West, Ma, & Aiken, 2012; Chin, Marcolin, & Newsted, 2003; Lin et al., 2010; Marsh et al., 2014; Maslowsky, Jager, & Hemken, 2015) and test 100, 200, and 500 sample size conditions. I found biased estimates for the interaction effects when sample size equals 100. Therefore, in the present study, I test two sample size conditions: 250 and 500.

Reliability, ρ. We manipulated the reliability, ρ, for both X and M to be either .70, .80, or .90. A Reliability of .70 represents 49% of the total variance being the true score variance and has been viewed as the acceptable lower boundary of reliability for group comparison in clinical research. Low reliability conditions (i.e., $\rho < .70$) were not considered in our simulation setting. The error variance of the exogenous variables were varied by the level of reliability values. The sum of the error variances for the three items for each latent factor was 3.85, 1.70, and 1.00, corresponding to .70, .80 and .90 reliability, respectively. We varied the error variances of the three items proportionally for ξ_X, ξ_M, and η_Y. The error variance of the first item covered 44% of the total error variances in each latent predictor, followed by 33% of the second item, and 23% of the third item. In other words, we manipulated the error variances as (1.29, 0.73, 0.51) for .70 reliability, (0.75, 0.56, 0.39) for .80 reliability and (0.33, 0.25, 0.17) for .90 reliability.

Mplus 7.11 (Muthén & Muthén, 1998-2013) was used to generate 500 data sets for each condition. Given that the data were generated at the item level (i.e., three items per latent factor), we computed the mean composite score for X and for M by averaging the corresponding items. Hence, we had three new observed composite scores; namely, the two observed composite variables X and M, and the corresponding product (or observed interaction effect) term XM. Double-centering strategy (Lin et al., 2010) was applied in the analyses. Therefore, X and M were first mean-centered; the product term XM was first computed using the mean-centered X and M and then mean- centered afterward. XM mean-centered afterward. The four reliability estimates (Cronbach's alpha, omega, Revelle's omega, and GLB) were computed for each dataset by using R packages ("MBESS") and ("psych") (McNeish, 2017). The item-level measurement error variance in each dataset were constraint by using sample-specific reliability values. Analyses with both the LMS and RAPI methods were conducted using Mplus 7.11. The annotated Mplus syntax for specifying the models with these three methods is presented in Appendix B.

Evaluation Criteria

Six criteria were selected to evaluate the accuracy, precision, and power of the interaction effect estimate Υ_{XM} between RAPI and LMS methods with four types of reliabilities. Each criterion was summarized from 500 replications for each simulation conditions.

Average bias. The average bias of each simulation condition, $B(\theta_c)$ was calculated as:

$$B(\theta_c) = R^{-1} \sum_{r=1}^{R} (\hat{\theta}_{rc} - \theta_c), \tag{11}$$

where $\hat{\theta}_{rc}$ denotes the parameter estimate for replication r in condition c, θ_c represents the population parameter for θ in condition c, and R indicates the total number of replications. In this study, θ_c is the true interaction effect γ_{XM} in condition c, which equals 0.2.

Standardized bias. In addition to comparing the average bias in its original magnitude, the average bias can be interpreted in terms of parameter standard errors. The standard error of each population parameter will be calculated from 500 replications. Thus, standardized bias $SB(\theta_c)$ was defined as:

$$SB(\theta_c) = \frac{B(\theta_c)}{SE_{\theta_c}}, \tag{12}$$

where SE_{θ_c} is the standard error of θ_c (Collines, Schafer, & Kam, 2001). The standardized bias of the latent interaction effect estimates was compared with the cutoff value of 0.40. An absolute value < 0.40 was regarded as acceptable (Collins, Shafer, & Kam, 2001).

Standard error (SE) ratio. The standard error ratio, $SER(\theta_c)$ is calculated by the following formula:

$$SER(\theta_c) = {SE_{\theta_c}}^{-1}[R^{-1}\sum_{r=1}^{R}\left(SE_{\hat{\theta}_{rc}}\right)], \tag{13}$$

where $SE_{\hat{\theta}_{rc}}$ indicates the standard error of parameter estimate for replication r in condition c. Hence, $SER(\theta_c)$ represents the ratio of the average estimated standard error from the sample to the empirical standard error (standard deviation of $\hat{\gamma}_{XM}$). The SE ratio was designed to evaluate the precision of the parameter estimators. Estimators with smaller SE ratio show less variability across simulation replications. The criterion for evaluating SE ratio is the same as evaluating relative SE bias (Hoogland and Boomsma, 1998). Hence, the absolute value of the SE ratio between 0.9 and 1.1 was considered acceptable.

Root mean square error (RMSE). The RMSE quantifies the sampling variability (i.e., the standard deviation) of the parameter estimates. The RMSE was calculated to evaluate both the accuracy and precision of the parameter estimations for the three methods. The smaller the RMSE values, the more accurate the parameter estimations were across the 500 replications. As to our knowledge, the criterion for making an adequate RMSE value has yet been developed. The RMSE values were used from a relative standpoint, in which the RMSE estimate for a certain method or reliability calculation was compared with others under the same simulation conditions.

95% confidence interval (CI) coverage rate. The 95% confidence interval coverage rate was calculated as:

$$R^{-1}(no.\, of\ replications\ where\ the\ CI\ contains\ \theta)$$

where R indicates the total number of replications and θ denotes the population parameter. For the 95% CI coverage, the Wald interval was obtained, with a coverage rate > 91% considered acceptable (Muthén & Muthén, 2002).

Power and Type I error rate. The statistical power for detecting the non-zero interaction effects were examined. Power estimates refers to the percentage of rejecting the null hypothesis when the interaction effect occur in the population data across replications. Conventionally, power above .80 is consider sufficient in the present study. In the present study, the power for testing the interaction effects for each method and reliability combinations was compared with the power from the true model. For sample size equal 250 conditions, the powers for testing $\gamma_{XM} = 0.2$ were .59, .71, and .86 for reliability equals .7, .8 , and .9, respectively. For sample size equal 500 conditions, the powers for testing $\gamma_{XM} = 0.2$ were .89, .93, and .99 for reliability equals .7, .8 , and .9, respectively.

Results

Four reliability estimates (Cronbach's alpha, omega total, Revelle's omega total, and greatest lower bound) were applied to adjust for the exogenous composites' measurement errors under the RAPI and the LMS methods while estimating interaction effects. The estimation comparisons are shown in terms of the biases (average and standardized) and standard error ratio in Table 3. The results of root mean square error (RMSE), 95% CI coverage rate and power are displayed in Table 4.

Convergence and Inadmissible Solutions

While using LMS methods with Revelle's omega total as the reliability estimates under samples size = 250 and reliability = .9 conditions, three (0.6%) out of 500 replications yielded non-converge results. These three cases were excluded from the subsequent analyses. All the other replications were converged without any inadmissible solutions across simulation conditions.

Average and Standardized Bias of Υ_{XM}

As shown in Table1, mean estimate of the interaction effects, using alpha, omega total, and GLB yielded biases ranged from 0.06 to 0.08, regardless of sample size, the amount of measurement errors, and the methods for estimating interaction effects. These biases resulted in 30% of 40% overestimation, compared to the true interaction effect of

0.2. On the other hand, applying Revelle's omega total resulted in biases ranged from 0.01 (5%) to 0.03 (15%) across simulation conditions. Note that as the amount of measurement errors decrease, the average biases increase.

The standardized bias increases as sample size increases. Consistent to the results of the average bias, the standardized biases for utilizing alpha, omega total, and GLB are all over the recommended criteria of 0.40, with a range of 0.51 to 1.04. Revelle's omega total yielded adequate standardized biases (<.0.40) for $\boldsymbol{\rho}$s equal to .7 and .8 conditions with a range of 0.11 to 0.32, regardless of samples size and the choice of interaction methods. In ρ = .9 conditions, the standardized biases were around 0.41 and 0.59 for samples size 250 and 500, respectively.

SE Ratio of $\boldsymbol{\Upsilon_{XM}}$

As shown in Table 3, the SE ratios were within the range of 0.9 and 1.1 across most of the simulation conditions for using alpha (ranged from 0.97 to 1.06), omega total (ranged from 0.97 to 1.07), and GLB (ranged from 0.97 to 1.05). The only exception occur under the high measurement errors (ρ = .7) and small sample size (n = 100) with the LMS method, where the SE ratios were 1.12, 1.13, and 1.11 for alpha, omega total, and GLB, respectively. All the SE ratios for conditions related to Revelle's omega total were within the range of 0.9 - 1.1.

Table 3

Mean Estimate, Relative Bias, and Standard Error Ratio of the Latent Interaction Effect from 500 Replications[a] *(γ_{XM}=0.2)*

	N	$\boldsymbol{\rho}$	Average Bias				Standardized Bias				Standard Error Ratio			
			Alpha	Omega	Revelle	GLB	Alpha	Omega	Revelle	GLB	Alpha	Omega	Revelle	GLB
RAPI	250	.70	0.07	0.08	0.01	0.07	**0.52**	**0.55**	0.12	**0.51**	0.98	0.97	0.97	0.97
		.80	0.07	0.07	0.02	0.07	**0.62**	**0.63**	0.22	**0.62**	0.98	0.97	0.97	0.97
		.90	0.07	0.07	0.03	0.07	**0.76**	**0.76**	**0.41**	**0.76**	0.98	0.98	0.98	0.98
	500	.70	0.06	0.07	0.01	0.06	**0.68**	**0.74**	0.16	**0.68**	0.97	0.97	0.96	0.97
		.80	0.06	0.07	0.02	0.07	**0.84**	**0.85**	0.30	**0.83**	0.96	0.96	0.96	0.96
		.90	0.06	0.07	0.03	0.06	**1.03**	**1.04**	**0.59**	**1.03**	0.97	0.97	0.97	0.97
LMS	250	.70	0.06	0.06	0.01	0.06	**0.48**	**0.51**	0.11	**0.47**	**1.12**	**1.13**	1.02	**1.11**
		.80	0.06	0.06	0.02	0.06	**0.59**	**0.60**	0.22	**0.58**	1.05	1.06	0.99	1.05
		.90	0.07	0.07	0.03	0.06	**0.74**	**0.75**	**0.42**	**0.74**	1.00	1.00	0.96	1.00
	500	.70	0.06	0.06	0.02	0.06	**0.69**	**0.74**	0.18	**0.68**	1.06	1.07	0.96	0.96
		.80	0.06	0.06	0.02	0.06	**0.83**	**0.85**	0.32	**0.83**	0.96	0.95	0.96	0.95
		.90	0.06	0.06	0.03	0.06	**1.02**	**1.03**	**0.59**	**1.02**	0.97	0.98	0.96	0.97

Note. RAPI = reliability-adjusted product-indicator method; LMS = latent moderated structural equations method; *N* = sample size; $\boldsymbol{\rho}$ **=** reliability estimate**;** Alpha = Cronbach's alpha; Omega = omega total; Revelle = Revelle's omega total; GLB = greatest lower bond.
[a] For N=250, $\boldsymbol{\rho}$ = .90, and LMS method with Revelle reliability condition, the number of replications is 497.
[b] Values exceeding the recommended cutoffs are in bold.

Table 4

Root Mean Square Error, 95% Confidence Interval (CI) Coverage Rate, and Power of the Latent Interaction Effect from 500 Replications[a] *(*γ_{XM}*=0.2)*

			95% CI Coverage (%)				Root Mean Square Error				Power				
	N	$\boldsymbol{\rho}$	Alpha	Omega	Revelle	GLB	Alpha	Omega	Revelle	GLB	True	Alpha	Omega	Revelle	GLB
RAPI	250	.70	94.2	94.4	95.2	94.4	0.16	0.16	0.11	0.15	0.59	0.55	0.55	0.58	0.55
		.80	91.6	91.6	95.4	91.4	0.13	0.14	0.09	0.14	0.71	0.69	0.69	0.71	0.69
		.90	**89.0**	**89.0**	92.6	**89.0**	0.11	0.11	0.09	0.11	0.86	0.86	0.86	0.85	0.86
	500	.70	**90.0**	**89.8**	92.6	**89.0**	0.11	0.12	0.07	0.11	0.89	0.87	0.87	0.88	0.87
		.80	**85.4**	**85.2**	92.2	**85.8**	0.10	0.10	0.07	0.10	0.93	0.94	0.94	0.94	0.94
		.90	**79.6**	**79.0**	**89.0**	**79.6**	0.09	0.09	0.06	0.09	0.99	0.99	0.99	0.99	0.99
LMS	250	.70	**90.6**	**90.6**	93.6	90.8	0.13	0.14	0.10	0.13	0.59	0.62	0.62	0.60	0.62
		.80	**90.4**	**90.0**	94.4	90.4	0.12	0.13	0.09	0.12	0.71	0.72	0.72	0.72	0.72
		.90	**88.8**	**88.0**	93.0	**88.8**	0.11	0.11	0.08	0.11	0.86	0.86	0.86	0.86	0.86
	500	.70	**87.8**	**87.0**	92.4	**88.0**	0.10	0.11	0.07	0.10	0.89	0.89	0.89	0.89	0.89
		.80	**83.6**	**84.0**	91.4	**83.8**	0.10	0.10	0.07	0.10	0.93	0.94	0.94	0.93	0.94
		.90	**77.0**	**76.8**	**89.6**	**77.0**	0.09	0.09	0.06	0.09	0.99	0.99	0.99	0.99	0.99

Note. *N* = sample size; $\boldsymbol{\rho}$ **=** reliability estimate**;** PM = Path model; RAPI = reliability-adjusted product-indicator method; LMS = latent moderated structural equations method. [a] For N=250, $\boldsymbol{\rho}$ = .90, and LMS method with Revelle reliability condition, the number of replications is 497.
[b] Values exceeding the recommended cutoffs are in bold

A SE ratio smaller than 1 indicates that the sample-estimated SE is, on average, smaller than the empirical standard error, and vice versa. Although most of the SE biases displayed in Table 3 were negligible, a trend of SE bias smaller than 1 occurred for the RAPI method and larger than 1 occurred for the LMS method was observed.

Coverage Rate of 95% CI of Υ_{XM}

As shown in Table 4, the coverage rate for the Revelle's omega total were adequate, with a range from 89.0% to 95.2% for the RAPI method, from 89.6% to 93.6% for the LMS method, regardless of sample size and the magnitude of measurement errors. The two below 91% coverage rates occurred when sample size equals 500 and $\boldsymbol{\rho}$ equals .90. As for the alpha, omega total, and GLB, the CI coverage rates were above 91% criteria for using the RAPI methods under the sample size equals 250 and $\boldsymbol{\rho}$ equals .70 and .80 conditions. However, the coverage rates were below 91% when sample size equals 500 regardless of the amount of measurement errors and the interaction effect methods. Overall, applying Revelle's omega total had the highest chance of identifying the true effect among the four reliabilities we compared.

RMSE in Estimating Υ_{XM}

Generally, the RMSE values decreased as sample size increased or measurement errors decreased. The RMSEs for the Revelle's omega total were smaller than that for the other three reliability estimates. For example, under the sample size = 250 and ρ = .7 conditions, the RMSE of the interaction effect estimates was 0.11 for applying Revelle's omega total, whereas the RMSEs were 0.15-0.16 for applying the other three reliabilities. Overall, the parameter estimates yielded from methods using the Revelle's omega total were the most precise and accurate (i.e., RMSE ranged from 0.06 to 0.11) among the four reliabilities. Finally, no obvious difference in RMSE was observed between the LMS and RAPI methods.

Power in Estimating Υ_{XM}

The power for testing the parameter estimates for each simulation condition was compared with the power of fitting the simulated data with true model. As shown in Table 4, the power of using the true model ranged from .59 to .99. Across all the simulation conditions, the powers were close to the power from the true model regardless of sample size, magnitude of the measurement errors, and interaction methods. Results indicate that using either the RAPI or LMS methods along with alpha, omega total, Revelle's omega total, or GLB reliabilities recovered the original power of the significant test.

Discussion

While estimating the latent interaction effects with manifest composites, the measurement error variance of the exogenous variables have to be constrained for model specification purpose. One way to apply the constraints is by using the reliability estimates of the scale to pre-calculate the measurement error variance of the exogenous manifest variables before running the models. Among the reliability estimates which can be applied in this scenario, Cronbach's alpha is easy to obtain in software and commonly reported. However, the item assumptions embedded within the usage of Cornbach's alph (e.g., tau-equivalent) are almost always violated. Other alternatives including omega total, Revelle's omega total, and greatest lower bond (glb) have been shown to provide more accurate reliability estimates. Therefore, in the present study, the four reliability estimates are utilized with both the RAPI and the LMS methods and the results in terms of the interaction effect estimation are compared.

In the comparisons among two different samples sizes (250 and 500) and true reliability values (.70, .80, and .90), the Revelle's omega total obtains unbiased and stable estimates. On the contrary, Cronbach's alpha, omega total, and glb provide biased and less stable estimates. Recall in study 1, omega total (or Cronbach's alpha) yielded unbiased interaction effects estimates with tau-equivalent items. Results from the present study have shown that when items are following congeneric equivalent, Cronbach's alpha, omega total, and glb are all underperformed. The four reliability estimates do not shown much difference in the power for detecting the interaction effects. All the powers for the four reliabilities are equal or close to the powers for the true models. Hence, if researchers are only interested in whether the interaction effects are significant or not, all the four reliability estimates are feasible to answer this question. However, the revelle's omega total is recommended if the real values of the effects are of interest.

Both the reliability and the sample size influence the interaction effects estimation. As the sample size increase, the standardized bias increase and the 95% CI coverage rate decrease. Given that the average bias stay the same regardless of samples, we can conclude that the change in both the standardized bias and the coverage rate may be due to the shrinkage of standard errors related to the sample size increase. Although the revelle's omega outperform the other three candidates, the interaction effects estimates provided by methods using revelle's omega increase biases as the true reliability of the scale increase. One possible explanation is that the *Schimd-Leiman rotation* applied to the revelle's omega calculation overly correct the measurement error variance and results in more biases when there is not much measurement errors to account for (i.e., $\boldsymbol{\rho}$ = .90).

A slight difference has been observed between the RAPI and the LMS methods. From the results of the SE ratio, all the values related to the RAPI methods are below 1 whereas most of the values related to the LMS methods are over 1, indicating that the interaction estimates from the LMS methods are less stable. Less stable interaction estimates through the LMS methods become salient (SE ratio > 1.1) when sample size equals 250 and the true reliability equals .70. Therefore, cautious should be made for researchers using the LMS methods with small sample size and low reliability items, especially with Crobach's alpha, omega total, and glb as the estimates to correct for the measurement errors of the exogenous composites.

Two limitations should be addressed. First, the item error variances were assumed independence in the present study. Independent error variance among items may not be true in the real world data and may result in biased reliability estimates (Green & Yang, 2009). Future research can investigate the impact of the correlation among item errors on the estimation of the interaction effects. Secondly, the issue of multicollinearity of the latent exogenous variables has not yet to be discussed in the present study. Since the estimation of the interaction effects is influenced by the exogenous variables, the correlation of the two exogenous variable may as well play a role in the interaction effects estimates. In the present study, I fixed the correlation in a small magnitude (r = .20). Future studies should be proceeded to investigate the impact of multicollinearity on the interaction effects estimates with composites.

CHAPTER IV

THE INTERPLAY BETWEEN INTERACTION EFFECTS AND QUADRATIC EFFECTS WITH COMPOSITES IN ADVANCED LATENT VARIABLE MODELS

Introduction

Testing nonlinear effects, including interaction and quadratic effects, has long been an important issue in social and behavior research. Interaction effects refer to the relationship between two variables stays the same or changes depending on the level of a third variable (i.e., the moderator). For example, educational researchers often hypothesized a particular teaching strategy interacts with students' characteristics, such as gender, literacy level, behavioral problem, etc., in determining learning outcomes. Quadratic effects indicate the association between the exogenous and endogenous variables steadily change to an optimal level, and then level off or even change oppositely beyond this optimal point. For instance, students may perceived higher teachers' teaching effectiveness as teachers assigned more assignments to them, given low to moderate demanded workload; however, the positive association between students' workload and teachers' teaching effectiveness may decrease for larger amount of assignments (Marsh, 2001). Interaction effects and nonlinear effects can be either solely exist or coexist, and may influence each other in the model (Ganzach, 1997; Busemeyer & Jones, 1983).

Structural equation modeling (SEM) has been a common way to estimate the nonlinear effects with unobserved variables. Within the SEM framework, the unobserved, or latent exogenous and endogenous variables are formulated in structural equations, and they are measured with measurement errors by observed indicator variables in a measurement model. Previous simulation studies have been mainly focus on interaction effects items (Jöreskog & Yang, 1996; Kenny & Judd, 1984; Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Lin, Wen, Marsh, & Lin, 2010; Little, Bovaird, & Widaman, 2006; Marsh, Wen, & Hau, 2004; Moulder & Algina, 2002; Wall & Amemiya, 2001), little attention has been given to quadratic effects. Additionally, methodologist generally focus on the methods be applied to exogenous variables measured by multiple items, little has been discuss for the observed manifests measured by composite scores.

In Chapter 2, I conduct a simulation study to evaluate methods for estimating latent interaction effects in latent variable models, and found that the latent moderated structure equation (LMS) method and the reliability-adjusted product indicator (RAPI) method outperform the conventional path analysis in terms of the estimation accuracy of the interaction effects, by modeling measurement errors of the exogenous variables. In Chapter 3, the simulation results show that the Revelle's omega total outperform the other three candidates (Cronbach's alpha, omega, and glb) and is recommended to while conducting the LMS and RAPI methods to estimate interaction effects with observed composites. As the pilot studies to address the issues of using composite scores in estimating latent interaction effects, some puzzles are remained to be solved. First, whether the study conclusions stand if the researchers change their focus from interaction effects to quadratic effects, is unknown. Harring, Weiss, and Hsu (2012) compared latent variable methods in estimating the quadratic effects but did not consider the used of composite variables. Secondly, several studies have found that the correlation between the linear effects has substantial influence on the estimation of the nonlinear effects (Kelava & Brandt, 2009, Kelava, Moosbrugger, et al., 2008, and Kelava, Werner, et al., 2011), but neither did they consider the condition of using manifest composites nor evaluate the robustness of the RAPI and LMS methods to multicollinearity.

Hence, the purpose of the present study is to expend the work from Chapter 2 and Chapter 3 by (1) taking more design factors, such as multicollinearity and (2) the occurrence of the quadratic effects, into account. Comparisons between the RAPI and the LMS methods and the conventional path analysis, which lack the assumption of measurement errors of the exogenous variables, will also be addressed. By conducting a series of Monte-Carlo simulation studies, how the LMS method, the RAPI method, and the conventional path analysis perform in terms of the precision and accuracy of the nonlinear effects estimations, will be investigated. Since the use of composites in estimating nonlinear effects is a common practice, it is important to examine the impact of ignoring measurement errors of the exogenous variables on nonlinear effects estimations.

Nonlinear Structural Equation Models

In social and behavioral research contexts, the relationships among variables has often been assumed to be linear. However, models carry with such linear relationship assumptions between variables often facing the challenge of not representing the reality in a better way (Kelava & Brandt, 2009). In particular, the relationship between an exogenous and an endogenous variable may (1) depends on a third variable, or (2) occur both linear and quadratic patterns. The former is called interaction effect and the latter is known as quadratic effect.

The structural model and the measurement model for latent interaction effects and latent quadratic effects are discussed below.

Latent Nonlinear Effect Model with Multiple Item

A typical structural equation model is composed by two regression models, which play distinctive roles in interpreting variables' association: (a) the structural model that defines the relationships among exogenous and endogenous variables, and (b) the measurement model that defines the relationships among latent and observed variables. Below I will describe the two models specified for estimating the latent nonlinear effects.

Structural model. When the effects of two latent variables (ξ_1 and ξ_2), their corresponding quadratic terms (ξ_1^2 and ξ_2^2), and their interaction ($\xi_1\xi_2$) on an endogenous variable y (y is a single observed variable) is considered, the following latent variable model has been used to estimate the nonlinear effects:

$$y = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \omega_{11}\xi_1^2 + \omega_{22}\xi_2^2 + \delta, \quad (1)$$

where α is the intercept, γ_1 and γ_2 represent the linear effects, ω_{12} represents the interaction effect, and δ is the disturbance of y.

The full nonlinear structural equation model can then be specified in the following matrix expression:

$$y = \alpha + \Gamma\xi + \xi'\Omega\xi + \delta. \tag{2}$$

In equation 2, the common factors ξ is defined as a 2 x 1 matrix:

$$\boldsymbol{\xi} = \binom{\xi_1}{\xi_2}. \tag{3}$$

Γ = (γ_1, γ_2) denotes the coefficient vector for the linear effects, $\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ 0 & \omega_{22} \end{pmatrix}$ is the upper triangular coefficient matrix of the nonlinear effects (with the quadratic effects on the diagonal and the interaction effects off-diagonal), and δ is the latent disturbance.

Measurement model. Assuming ξ_1 and ξ_2 each measured by 3 observed indicator variables, the measurement model for ξ_1 and ξ_2 can be described in a conventional confirmatory factor analysis (CFA) latent variable model (Bollen, 1989):

$$\boldsymbol{X} = \boldsymbol{\tau_X} + \boldsymbol{\Lambda_X}\boldsymbol{\xi} + \boldsymbol{\delta}, \tag{4}$$

where $\boldsymbol{X}$ are the exogenous observed indicator variables. $\boldsymbol{\xi}$ is a 2 x 1 vector of common factor scores. The common factors $\boldsymbol{\xi}$ are the commonality features explaining the correlations among the observed indicator variables. $\boldsymbol{\tau_X}$ is are 6 x 1 vector of latent measurement intercepts, which represent the expected $\boldsymbol{X}$ scores when $\boldsymbol{\xi} = \mathbf{0}.$ $\boldsymbol{\Lambda_X}$ is a 6 x 2 matrix of factor loadings on $\boldsymbol{X}$. The factor loadings $\boldsymbol{\Lambda_X}$ represent the level of the linear relationship between indicators and the observed variables. $\boldsymbol{\delta}$ are the corresponding random measurement error factors of the exogenous and endogenous variables, respectively. $\boldsymbol{\delta}$ represents the portion of the indicators not explained by the factors. In the common factor models, each of the elements are defined below.

For $\boldsymbol{X},$ the measured variables, the measurement intercepts, and the measurement error scores can be defined as a 2p x 1 vectors as,

$$X = \begin{vmatrix} X_1 \\ X_2 \\ \vdots \\ X_6 \end{vmatrix} \qquad \tau_X = \begin{vmatrix} \tau_{x_1} \\ \tau_{x_2} \\ \vdots \\ \tau_{x_6} \end{vmatrix} \qquad \delta = \begin{vmatrix} \delta_{x_1} \\ \delta_{x_2} \\ \vdots \\ \delta_{x_6} \end{vmatrix}. \tag{5}$$

The factor loadings $\boldsymbol{\Lambda_X}$ are defined as a 6 x 2 matrix as,

$$\boldsymbol{\Lambda_X} = \begin{vmatrix} \Lambda_{x_1} & 0 \\ \Lambda_{x_2} & 0 \\ \Lambda_{x_3} & 0 \\ 0 & \Lambda_{x_4} \\ 0 & \Lambda_{x_5} \\ 0 & \Lambda_{x_6} \end{vmatrix}, \tag{6}$$

Latent Nonlinear Effect Model with Composites

When researchers use the composites to represent the exogenous variables, the structural model is the same as depicted in Equation (1). However, researchers will find the measurement model hard to identify when using latent variable model with composites. By aggregating the item scores to represent each variable, the latent variable model will be composed by latent factors each with single indicator. Without any further constraint, such model with only one observed variable loaded on the latent nonlinear factor is not identifiable (Bollen, 1989; Kline, 2011). In literature, the parameters in such a nonlinear structural equation model can be estimated with two methods, which I describe in the following section.

Reliability-Adjusted Product Indicator (RAPI). Assuming one exogenous composites, $X = \Sigma x_i$ as the single indicator for the latent variable, ξ_X with the corresponding reliability, $\rho_{XX'}$ and variance, $Var(X)$. Because $\rho_{XX'}$ is the function of error variance, $Var(\delta_X)$ and latent score variance, $Var(\xi_X)$ (Bollen, 1989), the error variance of X, $Var(\delta_X)$ can be shown as:

$$Var(\delta_X) = (1 - \rho_{XX'})Var(X). \tag{7}$$

Therefore, the error variance of the composite can be estimated by using the reliability and composite variance. Similar idea can be applied to the nonlinear effect term. Bohrnstedt and Marwell (1978) and Busemeyer and Jones (1983) have derived the formula for estimating the error variance of the nonlinear terms with reliability of the composites. The two quadratic terms, X_1X_1 and X_2X_2, their corresponding error variances can be represent as:

$$Var(\delta_{11}) = 2(\rho_{X_1X_1}{}' Var(X_1)^2(1 - \rho_{X_1X_1}{}')) + (1 - \rho_{X_1X_1}{}')^2 Var(X_1)^2, \tag{8}$$

$$Var(\delta_{22}) = 2(\rho_{X_2X_2}{}' Var(X_2)^2(1 - \rho_{X_2X_2}{}')) + (1 - \rho_{X_2X_2}{}')^2 Var(X_2)^2. \tag{9}$$

And the error variance of the interaction term, X_1X_2 is denoted as:

$$Var(\delta_{12}) = \rho_{X_1X_1}{}' Var(X_1)(1 - \rho_{X_2X_2}{}')Var(X_2) + \rho_{X_2X_2}{}' Var(X_2)(1 - \rho_{X_1X_1}{}')Var(X_1) + (1 - \rho_{X_1X_1}{}')Var(X_1)(1 - \rho_{X_2X_2}{}')Var(X_2). \tag{10}$$

Once the error variances of the composites for each exogenous are calculated, they are all fixed by the estimated values in the latent variable model for estimating the nonlinear effects. This method has the advantage of being feasible to implement in the modern SEM software, such as LISREL (Jöreskog & Sörbom, 1996) and Mplus (Muthén & Muthén, 1998-2013).

Latent Moderated Structural Equations. Rather than directly creating the product indicator, Klein and Moosbrugger (2000) proposed the latent moderated structural equation (LMS) method, which can directly estimate the nonlinear latent variable with specific data distributional assumptions. The latent nonlinear effects, ξ_1^2, ξ_2^2, and ξ_{12} are estimated by utilizing the joint distribution of the latent linear effects, ξ_1

and ξ_2. Note that the error variances of the linear effects should be constrained by using Equation (7). The EM algorithm is used to compute maximum likelihood estimates of the parameters. The LMS method is implemented in Mplus (Muthén & Muthén, 1998-2013).

Method

Research Scenario

This study compared three methods of examining the nonlinear effects with observed composite scores to determine the estimation accuracy of the nonlinear effects. A Monte Carlo simulation study were conducted to evaluate the performance of the RAPI method, the LMS method, and the path analysis in estimating nonlinear effects. In general, the method consisted of generating data for a model testing one latent interaction and two latent quadratic effects on a single observed criterion variable (Figure 6). Both the latent true factors and latent unique factor are assumed to follow a standard normal distribution (i.e., mean equals to 0 and variance equals to 1.0). In the population model, each of the manifest exogenous variables were defined by three indicators. The multiple indicators were generated under congeneric item assumption and moderate reliability values (.7). The correlations between the two linear effects are manipulated as 0, .5, and .8. The linear effects are designed to explain 10% in total of the variance in the criterion variable, under the condition of zero correlation between ξ_1 and ξ_2 (Kelava, Werner et al., 2011). The nonlinear effects are designed to each explained 0% or 5% of the overall variance in the criterion variable.

The results were evaluated by examining the extent to which the parameter estimates recover the generating parameter values, as indicated by measures of bias and relative bias, and the variability of the parameter estimates in terms of 95% CI coverage rate, and by examining the standard error bias and the root mean square errors. Below I describe the choice of design factors and the evaluation criteria in detail.

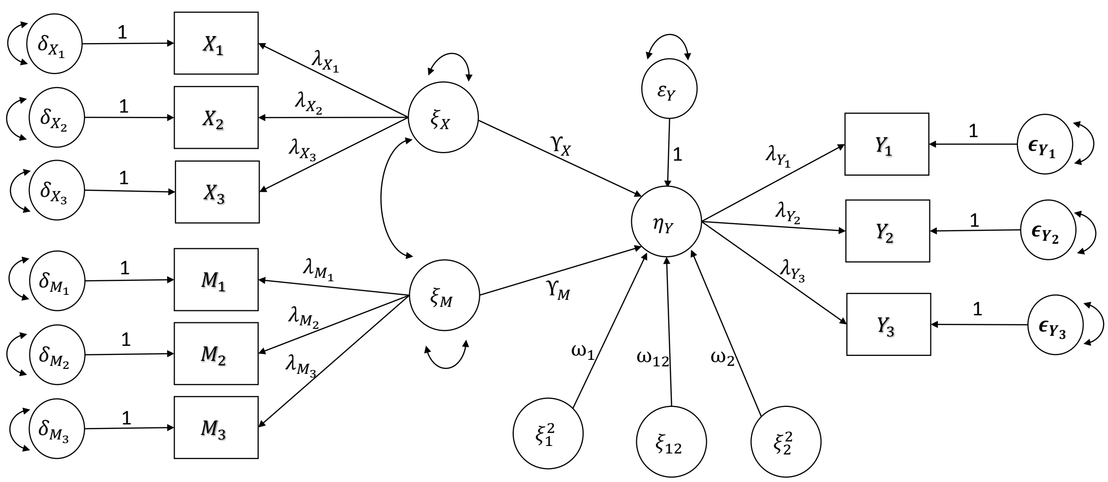

Figure 6. The pseudo population model with two latent linear effects and three latent nonlinear effects (one interaction and two quadratic effects). The nonlinear effects were created by using the latent moderate structural equations (LMS) method so no indicator has been created for the latent nonlinear effects.

Design of the Simulation Study

The manipulated variables were sample size, reliability obtained from the item scores within each exogenous variable, the reliability assumptions of the items wihin each exogenous variable, the effect size of the nonlinear effect, the level of multicollinearity between exogenous variables. In order to decide the conditions to be studies for each independent variable, previous studies with similar manipulations were considered.

Sample size. Previous studies on comparing different latent interaction testing approaches were used as references to decide the sample size conditions. Chin, Marcolin, and Newsted (2003) test interaction effect with partial least squares approach on sample size conditions of 20, 50, 100, 150, 200, and 500. Maslowsky, Jager, and Hemken (2015) proposed a simulation study to test the performance of the LMS approach with non-normal data on the sample size condition of 500. Cham, West, Ma, and Aiken (2012) estimating interaction with nonnormal observed data on sample size condition of 100, 200, 500, 1000, and 5000. Marsh, Wen, & Hau (2004) compared three strategies in estimating latent interaction under sample size conditions of 100, 200, and 500. Kelava, Werner et al (2011) studied how multicollinearity influence the estimation of nonlinear effect while both interaction and quadratic effects occurred in the model, with sample size equal to 400.

Given the five studies mentioned above, sample sizes of 100, 200, and 500 has been shown as commonly occurred levels in simulation studies. In the present simulation study, sample sizes of 500 were used to given small sample size have found to provide less precise estimates of the interaction effects in study 1.

Reliability values. Studies that examine the multicollinearity of the exogenous predictors on nonlinear effects estimations have included the condition with reliability value equal to .8 (Kelava, Werner, et al., 2011; Kelava, Moosbrugger, Dimitruk, and Schermelleh-Engel, 2008). Cham, West, Ma, and Aiken (2012) investigated the impact of data non-normality on the estimation of latent interaction effects, and the latent variable parameters fixed in their study result in .6 reliability, based on equation 2. Reliability value of .6 was also a condition assumed in Lin, Wen, Marsh, and Lin (2010), which compared double-mean centering and orthogonalizing strategies on the estimation of the latent interaction effect. Marsh, Wen, and Hau (2004) proposed .6 and .71 reliability values in their simulation studies for testing multiple latent interaction effect approaches.

Overall, the reliability values of .6, .7, and .8 has had been proposed in previous simulation studies. In the present study, the reliability, ρ, for both X and M were manipulated to be .70. A Reliability of .70 represents 49% of the total variance being the true score variance and has been viewed as the acceptable lower boundary of reliability for group comparison in clinical research. Low reliability conditions (i.e., $\rho < .70$) were not considered in our simulation setting.

Item factor loading and error variances. Under the SEM framework, both congeneric and tau-equivalent items (Lord & Novick, 1968) are allowed. In a CFA with zero latent means, tau-equivalent are defined as equal loadings but possibly unequal error variance across items; congeneric items are created by allowing both the factor loadings and error variances to vary for all measures (Jöreskog, 1971; Millsap & Everson, 1991).

Items' assumption in reliability has gain less attention in previous simulation studies on testing nonlinear effects. For the simulation studies conducted by Cham, West, Ma, and Aiken (2012) and Marsh, Wen and Hau (2004), the error variance of the exogenous indicators are assumed to be 1 and the factor loadings are allowed to vary in a certain level. Lin, Wen, Marsh, and Lin (2010) even assumed both factor loadings and error variances of the indicators are homogenous in item levels. Kelava, Werner, et al. (2011) and Kelava, Moosbrugger, Dimitruk, and Schermelleh-Engel (2008) considered the unequal parameters on both item factor loadings and error variance, but limited their discussions on only two levels of differences. For example, for each factor and the corresponding three items in each measurement model, the factor loadings and the error variances for the first item were fixed at 1.0 and .25, respectively; and the second and third items were both fixed at .894 and .20, respectively.

Since tau-equivalent items are hard to find in read data situation, we focus on the conditions with congeneric-equivalent items. Different values of population factor loadings for each indicator were randomly chosen from a uniform distribution within a range of 0.6 to 1 (rounded to two decimal places). The error variances for each of the three items are varied to explain 55%, 33%, and 12% of the variances in each latent predictor. As mentioned earlier, intercepts were fix at zero in the simulation study, as group comparisons is beyond the scope of the present study.

Multicollinearity. Multicollinearity refers to the high correlation between the exogenous variables in any model. In multiple regression, multicollinearity may result in parameter estimates with inflated standard errors and decreased power in detecting true effects. These become more serious issues in nonlinear latent model, given the unreliability of the indicators—the correlation between the latent factors is generally higher than the correlation between manifest indicators.

Despite the potential issue of multicollinearity, few studies have investigate the impact of multicollinearity on parameter estimation in nonlinear latent model, even the correlation between the exogenous latent predictors were manipulated. For example, in their simulation studies, Jaccard and Wan (1995) and Marsh et al. (2004) manipulated the multicollinearity levels of (.2, .4) and (.2, .3, .4) respectively. However, the simulation results were organized by aggregating the multicollinearity conditions. On the other hand, two past methodological studies has found multicollinearity of the latent predictors influence the performances of the methods for estimating nonlinear effects in SEM. In the two simulation studies conducted by Kelava, Moosbrugger et al (2008) and Kelave, Werner, et al (2011), the correlations between latent factors were manipulated as (0, 0.5, 0.8) and (0, .375, .625), respectively, and they found: as the multicollinearity increase, the distributional approach (LMS and QML) and the Jöreskog and Yang approach (1996) have the advantage in obtaining more accurate parameter estimates of the nonlinear effects.

For the present study conditions with 0, .5, and .8 correlation between latent predictors were examined. These numbers were chosen to cover the correlation scope designed in Jaccaed and Wan (1995) and Marsh et al (2004), and also be able to comparable to the conditions in the studies by Kelave et al (2008).

The effect size of the nonlinear effect. In the past methodologist studies, the magnitude of the nonlinear effect varied by how large the effect can explain the latent criterion's variance. Marsh, Wen, & Hau (2004) tested the performance of four different methods on estimating interaction effect, which explained 0%, 5% and 10% of the outcome variables' variance. Kelava, Moosbrugger et al. (2008) manipulated 0% and 5% effect size of the nonlinear effects to investigate the impact of multicollinearity and missingness on the latent nonlinear effect estimations. Kelava, Werner, et al. (2011) included conditions of 2.2% in testing the power of the nonlinear effects among the LMS, QML, and the unconstrained approaches.

In the present study, the magnitude of the nonlinear effects are set to explain 0% or 5% of the latent criterion variances. 0% effect size was selected to examine the Type I error rate when the null hypothesis is true (i.e., true nonlinear effects do not exist) (Cham, West, Ma, & Aiken, 2012). 5% effect size was chosen to investigate the power of detecting non-zero nonlinear effect. Three nonlinear models will be created base on the effect sizes of the quadratic and interaction effects: (1) interaction effects only (model 1), (2) quadratic effects only (model 2), and (3) both interaction and quadratic effects (model 3). Since the number of the nonlinear effects included in the analyses were different across the three population models, the corresponding path coefficients were differed among the three models. In model 1 and model 2, the nonlinear effects explained 0 % or 5% of the latent criterion's variance. In model 3, the interaction effect explained 0% or 5% of the latent criterion's variance, whereas the two quadratic effects were set to be equal in size and explained a total of 0% or 5% of the latent criterion's variance.

Software and Implementation

The data were generated in Mplus version 7.11 (Muthén & Muthén, 1998-2013) via Monte Carlo simulation procedure. Composite of item means and the mean-centered values for each exogenous variable were computed in R software (R core Team, 2013). The composites of the interaction or the quadratic terms were mean-centered and analyzed under conventional path analysis and models specified by using the RAPI and the LMS methods, using Mplus version 7.11. The different evaluation criteria measures, including standardized bias, standard errors of the parameter estimates, and the root mean square error were organized by using R software and Microsoft excel.

Evaluation Criteria

Six criteria were applied to evaluate the performance of the three methods in examining the nonlinear effect with observed composite scores.

95% confidence interval (CI) coverage rate. The 95% confidence interval coverage rate was calculated as:

$$R^{-1}(no. of\ replications\ where\ the\ CI\ contains\ \theta), \tag{11}$$

where R indicates the total number of replications and θ denotes the population parameter. For the 95% CI coverage, the Wald interval was obtained, with a coverage rate > 91% considered acceptable (Muthén & Muthén, 2002).

Raw bias. The raw bias of each simulation condition, $B(\theta_c)$ was calculated as:

$$B(\theta_c) = R^{-1} \sum_{r=1}^{R} (\hat{\theta}_{rc} - \theta_c), \qquad (12)$$

where $\hat{\theta}_{rc}$ denotes the parameter estimate for replication r in condition c, θ_c represents the population parameter for θ in condition c, and R indicates the total number of replications. In this study, when θ_c equals zero, the raw bias equals the mean of parameter estimates over 2000 replications.

Standardized bias. In addition to comparing the raw bias in its original magnitude, the raw bias will be interpreted in terms of parameter standard errors. The standard error of each population parameter will be calculated from 2000 replications. Thus, standardized bias $SB(\theta_c)$ was defined as:

$$SB(\theta_c) = \frac{B(\theta_c)}{SE_{\theta_c}}, \qquad (13)$$

where SE_{θ_c} is the standard error of θ_c (Collines, Schafer, & Kam, 2001; Merkle, 2011). The standardized bias of the latent interaction effect estimates was compared with the cutoff value of 0.40. An absolute value < 0.40 was regarded as acceptable (Collins, Shafer, & Kam, 2001).

Relative standard error (SE) bias. The relative SE bias, $SEB(\theta_c)$ is calculated by the following formula:

$$SEB(\theta_c) = {SE_{\theta_c}}^{-1}[R^{-1} \sum_{r=1}^{R} \left(SE_{\hat{\theta}_{rc}} - SE_{\theta_c}\right)], \tag{14}$$

where $SE_{\hat{\theta}_{rc}}$ indicates the standard error of parameter estimate for replication r in condition c. The relative SE bias was designed to evaluate the precision of the paramter estimators. Estimators with smaller relative SE bias show less variability across simulation replications. As recommended by Hoogland and Boomsma (1998), relative SE bias values < 10% were considered acceptable.

Root mean square error (RMSE). The RMSE quantifies the sampling variability (i.e., the standard deviation) of the parameter estimates. The RMSE was calculated to evaluate both the accuracy and precision of the parameter estimations for the three methods. The smaller the RMSE values, the more accurate the parameter estimations were across the 2,000 replications. A ratio of RMSE values for each design cell will be calculated to facilitate interpretation. For example, suppose the RMSE for the LMS method in a given design cell is 0.02, whereas the corresponding RMSE for an RAPI method is 0.03. The resulting ratio would be 1.50, suggesting that the RAPI estimates are, on average, 50% further away from the population nonlinear effect than the LMS estimates.

Power and Type I error rate. Both statistical power for detecting the non-zero nonlinear effects and Type I error of incorrectly detecting zero nonlinear effects as nonzero were examined. For conditions with zero nonlinear effects, Type I error rates refer to the percentage of retaining the null hypothesis of no nonlinear effect exist in the population data across replications. Type I error rate below .10 indicates acceptable. On the other hand, power estimates refers to the percentage of rejecting the null hypothesis when the nonlinear effect occur in the population data across replications. Power above .80 is consider sufficient in the present study.

Results

Three true models including true interaction model, true quadratic model, and the model with both true interaction and true quadratic effects, were specified. The model with one true interaction effect and two quadratic effects were used to analyze simulated data, regardless of the true models being used to generate the data. The two latent variable approach—the RAPI and the LMS methods—were compared in terms of estimating precision, accuracy, and their power to detect true effects across three levels of the multicollinearity between the two latent exogenous variables ξ_X and ξ_M. The true nonlinear effects along with the power for each true model are displayed in Table 5 through Table7.

Table 5

Results for Models with One True Interaction Effects

Cor.	Parameter	True value (Sig.)	Mean Est.	Bias%	Se ratio	95% CI	RMSE	Type I error	Power
				Reliability-Adjusted Product-Indicator					
0	ω_{12}	0.224 (91.6%)	0.249	2.23%	0.982	93.8%	0.073	N.A.	89.60%
	ω_{11}	0.000 (3.8%)	-0.002	N.A.	0.982	95.0%	0.040	5%	N.A.
	ω_{22}	0.000 (6.4%)	0.001	N.A.	0.906	91.6%	0.044	8%	N.A.
0.5	ω_{12}	0.200 (36.6%)	0.177	-11.55%	0.953	92.8%	0.089	N.A.	57.6%
	ω_{11}	0.000 (4.4%)	0.023	N.A.	1.009	93.4%	0.051	6.4%	N.A.
	ω_{22}	0.000 (6.4%)	0.025	N.A.	0.917	89.6%	0.056	10%	N.A.
0.8	ω_{12}	0.175 (13.0%)	0.121	-30.74%	0.937	89.8%	0.128	N.A.	22.2%
	ω_{11}	0.000 (6.6%)	0.037	N.A.	1.001	89.4%	0.071	10%	N.A.
	ω_{22}	0.000 (7.4%)	0.039	N.A.	0.923	87.0%	0.076	12.8%	N.A.
				Latent Moderate Structural Equations					
0	ω_{12}	0.224 (91.6%)	0.248	1.76%	0.990	94.8%	0.071	N.A.	90.4%
	ω_{11}	0.000 (3.8%)	-0.0018	N.A.	0.987	94.4%	0.049	5.6%	N.A.
	ω_{22}	0.000 (6.4%)	0.0013	N.A.	0.910	91.8%	0.054	7.8%	N.A.
0.5	ω_{12}	0.200 (36.6%)	0.191	-4.46%	0.946	92.4%	0.111	N.A.	44.6%
	ω_{11}	0.000 (4.4%)	0.014	N.A.	1.008	93.6%	0.066	6.4%	N.A.
	ω_{22}	0.000 (6.4%)	0.016	N.A.	0.914	90.4%	0.073	9.4%	N.A.
0.8	ω_{12}	0.175 (13.0%)	0.139	-20.47%	0.928	91.4%	0.261	N.A.	12.2%
	ω_{11}	0.000 (6.6%)	0.028	N.A.	0.975	93.0%	0.139	6.8%	N.A.
	ω_{22}	0.000 (7.4%)	0.031	N.A.	0.918	90.6%	0.147	9.4%	N.A.

Table 6

Results for Models with Two True Quadratic Effects

Cor.	Parameter	True value (Sig.)	Mean Est.	Bias%	Se ratio	95% CI	RMSE	Type I error	Power
				Reliability-Adjusted Product-Indicator					
0	ω_{12}	0.000 (5.0%)	-0.002	N.A.	0.957	93.8%	0.073	5.8%	N.A.
	ω_{11}	0.158 (88.8%)	0.129	-18.04%	0.989	87.4%	0.050	N.A.	90.0%
	ω_{22}	0.158 (87.6%)	0.132	-16.36%	0.932	88.0%	0.051	N.A.	88.8%
0.5	ω_{12}	0.000 (6.2%)	0.089	N.A.	0.947	78.6%	0.124	21.2%	N.A.
	ω_{11}	0.158 (55.0%)	0.119	-24.89%	0.998	85.6%	0.061	N.A.	73.2%
	ω_{22}	0.158 (56.0%)	0.121	-37.54%	0.924	84.6%	0.063	N.A.	71.4%
0.8	ω_{12}	0.000 (9.2%)	0.153	N.A.	0.938	70.2%	0.193	29.4%	N.A.
	ω_{11}	0.158 (12.0%)	0.099	-37.54%	0.996	82.8%	0.085	N.A.	37.4%
	ω_{22}	0.158 (15.0%)	0.101	-36.38%	0.918	79.2%	0.088	N.A.	39.2%
				Latent Moderate Structural Equations					
0	ω_{12}	0.000 (5.0%)	-0.000	N.A.	0.975	93.4%	0.072	6.6%	N.A.
	ω_{11}	0.158 (88.8%)	0.160	1.45%	0.993	94.0%	0.050	N.A.	89.0%
	ω_{22}	0.158 (87.6%)	0.163	3.43%	0.946	93.2%	0.054	N.A.	87.8%
0.5	ω_{12}	0.000 (6.2%)	0.046	N.A.	0.953	90.0%	0.121	10.0%	N.A.
	ω_{11}	0.158 (55.0%)	0.151	-4.23%	1.000	94.6%	0.068	N.A.	63.2%
	ω_{22}	0.158 (56.0%)	0.154	-2.78%	0.937	93.6%	0.072	N.A.	61.6%
0.8	ω_{12}	0.000 (9.2%)	0.106	N.A.	0.935	89.0%	0.283	10.8%	N.A.
	ω_{11}	0.158 (12.0%)	0.127	-19.84%	0.975	92.0%	0.143	N.A.	17.8%
	ω_{22}	0.158 (15.0%)	0.129	-18.63%	0.928	93.4%	0.149	N.A.	19.8%

Table 7

Results for Models with One True Interaction and Two True Quadratic Effects

Cor.	Parameter	True value (Sig.)	Mean Est.	Bias%	Se ratio	95% CI	RMSE	Type I error	Power
				Reliability-Adjusted Product-Indicator					
0	ω_{12}	0.224 (87.2%)	0.229	2.27%	0.950	94.0%	0.076	N.A.	87.6%
	ω_{11}	0.158 (86.0%)	0.130	-17.86%	0.966	87.4%	0.051	N.A.	89.0%
	ω_{22}	0.158 (86.2%)	0.132	-16.23%	0.902	85.8%	0.052	N.A.	86.4%
0.5	ω_{12}	0.200 (35.8%)	0.268	34.02%	0.940	86.0%	0.113	N.A.	87.2%
	ω_{11}	0.158 (54.6%)	0.143	-9.16%	0.985	91.4%	0.051	N.A.	85.6%
	ω_{22}	0.158 (54.2%)	0.145	-8.07%	0.899	91.2%	0.055	N.A.	81.8%
0.8	ω_{12}	0.175 (11.4%)	0.277	58.03%	0.931	84.4%	0.l59	N.A.	66.8%
	ω_{11}	0.158 (11.8%)	0.138	-12.96%	0.988	91.6%	0.067	N.A.	61.2%
	ω_{22}	0.158 (15.2%)	0.139	-12.03%	0.899	90.8%	0.072	N.A.	58.0%
				Latent Moderate Structural Equations					
0	ω_{12}	0.224 (87.2%)	0.232	3.68%	0.974	94.6%	0.075	N.A.	88.0%
	ω_{11}	0.158 (86.0%)	0.162	2.37%	0.981	94.0%	0.052	N.A.	87.2%
	ω_{22}	0.158 (86.2%)	0.165	4.66%	0.931	93.6%	0.056	N.A.	86.8%
0.5	ω_{12}	0.200 (35.8%)	0.240	20.23%	0.952	91.6%	0.122	N.A.	60.6%
	ω_{11}	0.158 (54.6%)	0.168	6.24%	1.004	94.2%	0.070	N.A.	68.2%
	ω_{22}	0.158 (54.2%)	0.170	7.63%	0.928	92.4%	0.076	N.A.	66.2%
0.8	ω_{12}	0.175 (11.4%)	0.246	40.85%	0.936	91.2%	0.277	N.A.	21.2%
	ω_{11}	0.158 (11.8%)	0.158	0.27%	0.978	94.6%	0.144	N.A.	24.0%
	ω_{22}	0.158 (15.2%)	0.160	1.38%	0.927	93.2%	0.152	N.A.	25.2%

True Interaction Model

The parameter estimates for the true interaction model are displayed in Table 5. As can be seen, when the correlation between the latent factors is zero, both approach resulted in unbiased parameter estimates. The relative biases for estimating interaction effects were below 3% and the quadratic effects estimates did not deviate from true values (0) by more than .002 in absolute value for both methods. In all cases, the SE ratios were within the range of 0.9 and 1.1 and the CI coverage rates were all above 91%. The RMSE for detecting the true interaction effects were around 0.07, higher than that for estimating the two null quadratic effects at around .045. The power in detecting the interaction effects was 89.6% and 90.4% for the RAPI and the LMS methods, respectively, which were lower than the power of the true model by 1% to 2%. Compared to the type I error rates from the true models (3.8% and 6.4%), slighted higher Type I error rates were observed for the first quadratic effect (5%) and the second quadratic effect (8%).

Given a correlation of .50 between the two latent exogenous variables, biased parameter estimates were found for the RAPI method with underestimated interaction effects of 11.55%. The LMS method as well underestimated the interaction effects but was within the criteria of 10%. In all cases, the SE ratios were within the range of 0.9 and 1.1. The CI coverage rates were all above 91% for the interaction effect and the first quadratic effect. The CI coverage rates for the second quadratic effect were 89.6% and 90.4% for the RAPI and LMS methods, respectively. The RMSE values for testing the interaction effects were 0.089 for the RAPI method and 0.111 for the LMS method.

Compared to the true model, the power for detecting the true interaction effects increased 20% and 8% for the RAPI and the LMS methods, respectively. The type I error rates for the two quadratic effects were higher at 2% to 3% from that in the true model.

When the correlation between ξ_X and ξ_M increased to .80, both methods yielded seriously underestimated the interaction effects. The relatively bias for the RAPI method was -30.74% and for the LMS method was -20.47%. The SE ratios for both methods were within the range of 0.9 and 1.1, indicating the biased estimates were stable. The 95% CI coverage were below the 91% criteria for the RAPI methods in estimating all the nonlinear effects, whereas the coverage rates for the LMS method were above 91% criteria for the interaction effect and the first quadratic effect. The RMSE values for estimating the nonlinear with the LMS methods were two times higher than that with the RAPI methods. Finally, both the power for detecting the interaction effects and the type I error rates for estimating the null quadratic effects were similar to or higher than the true model.

True Quadratic Model

Under the conditions of independent latent exogenous variables, substantially underestimated mean quadratic effect estimates were yielded through the RAPI method (Bias% = 18.04% and -16.36%). On the other hand, unbiased mean quadratic estimates were obtained using the LMS method (Bias% = 1.45% and 3.43%). The SE ratio for both methods were above within the .9 to 1.1 criteria but the CI coverage rate for the quadratic effects were below 91% for the RAPI method. The RMSE for estimating the interaction effects was around 0.07 and for examining the two quadratic effects were around .05 regardless of the methods. Slightly higher Type I error rate for the null interaction effect and power for the non-zero quadratic effects were observed for both methods.

When the correlation between ξ_X and ξ_M increased to .50, the quadratic effect estimates for the RAPI method yielded larger biases, with relatively bias of -24.89% and -37.54%. On the contrary, unbiased mean parameter estimates for the quadratic effects were obtained, with relatively bias of -4.23% and -2.78%). The SE ratio for both methods were within the range of 0.9 and 1.1. The 95% CI coverage rates for the LMS method were over 91% for all the three nonlinear effects estimates, whereas 78.6% of the CI for the interaction effects and 85% of the CI for the quadratic effects produced through the RAPI method can capture the true effects. The RMSE for detecting the (nonsexist) interaction effect increased to 0.120 for both methods. The type I error rate obtained from the RAPI method was three times larger than that from the true model (21.2% vs. 6.2%); an inflated Type I error rate was also observed from the LMS method but it was closer to the true model (10.0% vs. 6.2%). Compared to the power for detecting the true quadratic effects in the true model (55.0% and 56.0%), both methods produced higher power test on the quadratic effects, with the RAPI yielded 15% higher and the LMS method yielded 6% higher in power.

As the correlation between the two latent exogenous variables increased to .80, both the RAPI and the LMS obtained substantially biased mean estimates of the quadratic effects, with two times larger biases yielded from the RAPI method. The 95%

CI coverage rate of the nonlinear effects estimates were slightly decreased to 89.0% to 93.4% for the LMS method, whereas the rate were decreased to 70.2% to 82.8% for the RAPI method. The type I error rate was once again more conservative for the LMS method (10.8%) than the RAPI method (29.4%). Likewise, the power was higher for the RAPI method than the LMS method.

True Model with Interaction and Quadratic Effects

Similar to what we have seen in the previous two models, as long as the latent exogenous variables are uncorrelated, unbiased mean parameter estimates of the nonlinear effects were obtained for the LMS methods when a model with both effect types was estimated (Table XX). On the other hand, the quadratic effects yielded from the RAPI methods were substantially underestimated, whereas the interaction effect was unbiased. The power for detecting nonlinear effects were around 87%, which were similar to the power from the true model.

When the correlation between the two latent exogenous variables increased to .50, the relative biases of the interaction effects were 34.02% for the RAPI and 20.23% for the LMS methods. Likewise, the RMSE for estimating the interaction effects increased to 0.113 for the RAPI and 0.122 for the LMS. For the quadratic effects, underestimated estimates were observed from the RAPI method whereas the estimates in the LMS method were overly estimated, but all of them were below the 10% criteria in absolute value. In terms of power, the RAPI method maintain the power around 85% whereas the LMS method had power around 65%, both were higher than the power form the true model (35.8% to 54.6%).

As the correlation between ξ_X and ξ_M increased to .80, the biases of the interaction effects increased to 58.03% for the RAPI method, and 40.85% for the LMS method. The quadratic effects for the RAPI methods were slightly over 10% criteria in absolute value, whereas the quadratic effects for the LMS methods were unbiased. Although higher biases were observed for the RAPI methods, the RMSE of the nonlinear effects for the RAPI methods were half the size of that for the LMS methods, indicating more variation among LMS parameter estimates. Both method provided higher power test than the true model, with the power for the RAPI method higher than the LMS method by 40%.

Discussion

The primary goal of this article is to investigate the robustness of the RAPI and the LMS methods to multicollinearity while estimating nonlinear effects (both interaction and quadratic effects) with composite scores. Our results reveal that each method perform differently in terms of the estimation of the interaction effects or quadratic effects, and differ from the robustness to multicollinearity.

When the latent exogenous variables are uncorrelated, both the RAPI and the LMS methods provide unbiased estimates and have sufficient power for testing interaction effects. However, for estimating quadratic effects, the LMS method outperforms the RAPI method in providing unbiased estimates. Specifically, specifying model with both interaction effects and quadratic effects leads to unbiased estimates with the LMS method when latent exogenous variables are uncorrelated. The results are consistent with Keleva et al., (2008), in which LMS method are used for testing nonlinear effects with multiple predictors and are found to be unbiased.

Under increasing multicollinearity, both methods provide more biased estimates of the nonlinear effects, when the true models are only carrying interaction effects or quadratic effects. In all cases, the LMS method also has shown more accurate estimates of the nonlinear effects than the RAPI method. However, for true models with both interaction effects and quadratic effects, the estimation of the quadratic effects are unbiased in the LMS method and showing more robustness in the RAPI method, regardless of the level of multicollinearity. On the other hand, the interaction effects estimates are overestimated in a true model with both nonlinear effects, as opposed to underestimated in a true model with only interaction effect. Hence, specifying both interaction effects and quadratic effects in a model can be beneficial for examining quadratic effects, but may not be so helpful for investigating interaction effects.

The results of non-robustness of the interaction effects are in contradict to Keleva et al. (2008), in which they found that the LMS method is robust to multicollinearity among the three true models. Compared to Keleva et al. (2008) in which the latent variables are specified with multiple predictors, the measurement error adjustment approach in the present study seems to be “over-killed” and result in less accurate estimates. However, both methods provide higher power test of the nonlinear effects regardless of the level of multicollinearity. For researchers only interested in whether the models are able to detect the nonlinear effects, but not the exact magnitude of the effects, both the RAPI and the LMS methods may be preferable.

Combine with the simulation setting in study 2, we can conclude that unbiased estimates of the interaction effects are obtained when the correlation between the two latent exogenous variables is within the range of 0 to 0.2. When specifying both interaction effects and quadratic effects in a model, the quadratic effects estimates are more robust to multicollinearity than the interaction effects. Hence, it is recommended to use models specifying both interaction effects and quadratic effects, when researchers are only interested in quadratic effects. Overall, LMS method is better than the RAPI method in all cases in this study. However, both methods are suffered from the multicollinearity of the latent exogenous variables. In sum, when testing nonlinear effects with observed composites, both the RAPI and LMS methods are not recommended with moderate or higher correlations between latent exogenous variables, unless only the quadratic effects are of interest.

CHAPTER V

CONCLUSIONS

The dissertation discuss the latent variable models and statistical methods which can be applied while estimating nonlinear effects (i.e., interaction effects and quadratic effects) with manifest composites. As composite scores have been widely used among substantive research and the latent factor components (i.e., true score and measurement errors) have commonly been overlooked, much more work is needed to address the importance of using latent variable models while estimating nonlinear effects.

In the first manuscript of this dissertation, I evaluated two methods— reliability adjusted product indicator (RAPI) and latent moderate structural equations (LMS)— which can be applied while estimating interaction effects with composite scores. The two methods were also compared with the conventional path analyses, in which the exogenous variables are all assumed to be perfectly measured. I found that when examining an interaction effect based on the observed composite scores without properly taking measurement errors into account, the result may be a considerable underestimation in the interaction effect. Thus, we encourage researchers to apply either the LMS or the RAPI method, which can directly specify the measurement errors of the manifest variables, for the estimation of interaction effects. For researchers who have very limited access to SEM programs, the RAPI model is by far the most feasible way (i.e., can be implemented in most of the SEM programs) to generate unbiased interaction estimates. Moreover, the overall model chi-square test and other commonly used model-fit indices are only available for the RAPI method. On the other hand, the LMS method produces relatively more conservative interaction effect estimates. Additionally, for those who have small data sets (with low sample sizes) or less reliable measures, the LMS method would be more preferable.

One key feature of the RAPI and the LMS methods is to utilize the scale reliability to constrain the measurement error variance of the composite variables in the model. In the second manuscript, I evaluated the performance of both the RAPI and LMS methods with four different reliability estimates (i.e., Cronbach's alpha, omega, Revelle's omega, and greatest lower bond). The simulation results showed that incorporating with different reliability estimates would not substantially vary the power for testing the interaction effects. However, if the estimation accuracy and precision are of interest, Revelle's omega outperform the other three reliability estimates and is recommended when items followed the congeneric assumptions.

In the third manuscript, the focus is no longer only on the interaction effects, but is extended to quadratic effects as well. Results from the simulation showed that generally the LMS perform better than the RAPI methods while estimating for both interaction effects and quadratic effects in the model. However, when the correlation of the two latent exogenous variables are high (i.e., over .5), both the LMS and the RAPI methods are likely to yield biased estimations of the nonlinear effects. Additionally, the generated data were fitted with models assuming both interaction effects and quadratic effects occurred. Such models may increase the estimation accuracy while the researchers are only interested in quadratic effects, but may result in biased estimation when interaction effects are the only interest. Hence, prior knowledge which provides theoretical support in the model specification process is important, given that fitting models with both nonlinear effects may not always yielded ideal estimation of the nonlinear effects.

The findings of the three manuscripts in this dissertation can be summarized in the following recommendations.

1. While estimating nonlinear effects with observed composites, always conduct latent variable models and apply both the RAPI and LMS methods.
2. For items following congeneric assumption, applying the RAPI and LMS methods with the Revelle's omega total yielded more accurate results; if only the power of the test is of interest, applying Cronbach's alpha, omega, and GLB make less difference.
3. Caution should be made for applying the RAPI and LMS methods when the correlation between the latent exogenous variables are high (i.e., over .5).
4. To achieve better estimation, apply interaction effect models while only interested in estimating interaction effects; apply models with both interaction effects and quadratic effects while interested in estimating quadratic effects only.

REFERENCES

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

Bentler, P., & Woodward, J. (1980). Inequalities among lower bounds to reliability: With applications to test construction and factor analysis. *Psychometrika*, *45*, 249-267.

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

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.

Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. New York, NY: Cambridge University Press.

Busemeyer, J. R., & Jones, L. E. (1983). Analyses of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin*, *93*, 549-562.

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

Chin, W. W., Marcolin, B. L., Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a

Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, *14*, 189-217.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. London, UK: Lawrence Erlbaum Associates.

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*, 300-315.

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6,* 330-351.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: Holt, Rinehart & Winston.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of a test. *Psychometrika*, *16*, 297-334.

Dunlap, W. P., & Kemery, E. R. (1988). Effects of predictor intercorrelations and reliabilities on moderated multiple regression. *Organizational Behavioral and Human Decision Processes*, *41*, 248-258.

Eby, L. T., Butts, M. M., Hoffman, B. J., & Sauer, J. B. (2015). Cross-lagged relations between mentoring received from supervisors and employee OCBs:

Disentangling causal direction and identifying boundary conditions. *Journal of Applied Psychology*, *4*, 1275-1285.

Evans, M. G. (1985). A Monte Carlo study of the effects of correlated method variance in moderated multiple regression analysis. *Organizational Behavioral and Human Decision Processes*, *36*, 305-323.

Feucht, T. E. (1989). Estimating multiplicative regression terms in the presence of measurement error. *Sociological Methods and Research*, *17*, 257-282.

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

Green, S. B., & Yang, Y. (2009). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika, 74*, 155-167.

Harring, J. R., Weiss, B. A., & Hsu, J. –C. (2012). A comparison of methods for estimating quadratic effects in nonlinear structural equation models. *Psychological Methods*, *17*, 193-214.

Hattie, J. (1985). Methodology review: assessing unidimensionality of tests and items. *Applied Psychological Measurement*, *9*, 139-164.

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

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.

Jackson, P. H., & Agunwamba, C. C. (1977). Lower bounds for the reliability of the total score on a test composed of non-homogeneous items: I: Algebraic lower bounds. *Psychometrika*, *42*, 567-578.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric test. *Psychometrika*, *36*, 109-133.

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.

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

Kelava, A., & Brandt, H. (2009). Estimation of nonlinear latent structural equation models using the extended unconstrained approach. *Review of Psychology*, *16*, 123-131.

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

Kelava, A., Werner, C. S., Schermelleh-Engel, K., Moosbrugger, H., Zapf, S., 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,* 465-491.

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

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

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.

Kline, R. B. (2011). *Principles and practice of structural equation modeling.* New York, NY: Guilford Press.

Kwok, O., Im, M., Hughes, J. N., Wehrly, S. E., & West, S. G. (2016). Testing statistical moderation in research on home-school partnerships: Establishing the boundary conditions. In S. M. Sheridan & E. M. Kim (Eds.), *Research on family-school partnerships: An interdisciplinary examination of state of the science and critical needs (Volume III: Family-school partnerships in context)* (pp.79-107). New York, NY: Springer.

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

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

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

McDonald, R. P. (1978). Generalizability in factorable domains: "domain validity and generalizability". *Educational and Psychological Measurement*, *38*, 75-79.

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

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

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*, 87-96.

McNeish, D. (2017). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*. doi: 10.1037/met0000144

Millsap, R. E., & Everson, H. (1991). Confirmatory measurement model comparisons using latent means. *Multivariate Behavioral Research*, *26*, 479-497.

Mitchell, M. S., Vodel, R. M., & Folger, R. (2015). Third parties' reactions to the abusive supervision of coworkers, *Journal of Applied Psychology, 4,* 1040-1055.

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

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, *9*, 599-620.

Muthén, L. K., & Muthén, B. O. (1998-2013). *Mplus user's guide*. 7th ed. Los Angeles, CA: Muthen & Muthen.

Qin, X., Ren, R., Zhang, Z.-X., & Russell, E. J. (2015). Fairness heuristics and substitutability effects: Inferring the fairness of outcomes, procedures, and interpersonal treatment when employees lack clear information. *Journal of Applied Psychology*, *3*, 749-766.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, *21*, 173-184.

Raykov, T., & Shrout, P. E. (2002). Reliability of scales with general structure: Point and interval estimation using a structural equation modeling approach. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*, 195-212.

Revelle, W. (2016). *Using R and the psych package to find ω*. Retrieved from http://personality-project.org/r/psych/HowTo/omega.pdf.

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, *74*, 145-154.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*, 53-61.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*, 107-120.

Yang, Y. & Green, S. B. (2010). A note on structural equation modeling estimates of reliability. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*, 66-81.

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

APPENDIX A

ERROR VARIANCE OF THE LATENT INTERACTION EFFECT

The following is a summary of the derivation based on Bohrnstedt and Marwell (1978) and Busemeyer and Jones (1983). Let X (predictor) and M (moderator) be observable random variables with true scores ξ_X and ξ_M and error random variables δ_X and δ_M. We assume the following measurement models for X and M, respectively:

$$X = \tau_X + \lambda_X \xi_X + \delta_X, \tag{A1}$$

$$M = \tau_M + \lambda_M \xi_M + \delta_M. \tag{A2}$$

Both X and M are mean-centered variables so that $E(X) = E(M) = 0$. For identification purpose, both $E(\xi_X)$ and $E(\xi_M)$ are fixed to zero. Thus, the two intercepts, τ_X and τ_M, would be equal to zero. λ_X and λ_M are factor loadings that are constrained to one for identification purpose; these constraints allow the observed variables and the true scores to share the same metric. δ_X and δ_M are assumed to be independent from each other as well as independent from ξ_x and ξ_M, with $E(\delta_X) = E(\delta_M) = 0$. The variance of ξ_X is defined as:

$$Var(\xi_X) = E({\xi_X}^2) - \left(E(\xi_X)\right)^2 = E({\xi_X}^2), \tag{A3}$$

and the variances of ξ_M, δ_X, and δ_M can all be, respectively, found using the definition in Equation (A3): $Var(\xi_M) = E({\xi_M}^2)$, $Var(\delta_X) = E({\delta_X}^2)$, and $Var(\delta_M) = E({\delta_M}^2)$.

The observed interaction variable, XM, is defined as the product term of the two observed composite variables X and M. The corresponding latent true score of XM, ξ_{XM} is defined as the product term of ξ_X and ξ_M, so $\xi_{XM} = \xi_X \xi_M$. As Lin and colleagues (2010) pointed out, the use of double-mean-centering strategy can produce more accurate results when estimating latent interaction effect. Therefore, we adopted the double-mean-centering strategy; XM is also a mean-centered variable. The variance of this observed interaction variable XM is defined as:

$$Var(XM) = E(X^2M^2) - \left(E(XM)\right)^2, \tag{A4}$$

in which,

$$\begin{aligned} E(X^2M^2) &= E((\xi_X + \delta_X)^2(\xi_M + \delta_M)^2) \\ &= E\left(\left({\xi_X}^2 + 2\xi_X\delta_X + {\delta_X}^2\right)\left({\xi_M}^2 + 2\xi_M\delta_M + {\delta_M}^2\right)\right) \\ &= E\left({\xi_X}^2{\xi_M}^2\right) + E\left({\xi_X}^2{\delta_M}^2\right) + E\left({\delta_X}^2{\xi_M}^2\right) + E\left({\delta_X}^2{\delta_M}^2\right) \\ &\quad +2E(\delta_M)E\left({\xi_X}^2\xi_M\right) + 2E(\delta_X)E(\xi_X{\xi_M}^2) + 2E(\delta_X)E\left(\xi_X{\delta_M}^2\right) \\ &\qquad + 2E(\delta_M)E\left(\xi_M{\delta_X}^2\right) \\ &\quad +4E(\delta_X)E(\xi_X\xi_M\delta_M) \\ &= E\left({\xi_{XM}}^2\right) + E\left({\xi_X}^2\right)E({\delta_M}^2) + E\left({\delta_X}^2\right)E({\xi_M}^2) + E\left({\delta_X}^2\right)E({\delta_M}^2) + 0 + 0 + \\ &\quad 0 + 0 + 0, \end{aligned} \tag{A5}$$

and

$$\left(E(XM)\right)^2 = (E((\xi_X + \delta_X)(\xi_M + \delta_M)))^2$$
$$= \left(E(\xi_X\xi_M) + E(\xi_M\delta_X) + E(\xi_X\delta_M) + E(\delta_X\delta_M)\right)^2$$
$$= (E(\xi_{XM}) + 0 + 0 + 0)^2. \quad \text{(A6)}$$

In Bohrnstedt and Marwell (1978) and Busemeyer and Jones (1983), the derivations of both Equations (A5) and (A6) are based on the assumptions of bivariate normality in X and M. However, when applying the double-mean-centering strategy (Lin et al., 2010), Equations (A5) and (A6) may be derived without any distribution assumption on X and M (other than the assumption that the variances of ξ_X, ξ_M, δ_X, and δ_M are finite). When substituting Equations (A5) and (A6) back into Equation (A4), we get

$$Var(XM) = [E\left(\xi_{XM}{}^2\right) - \left(E(\xi_{XM})\right)^2] + E\left(\xi_X{}^2\right)E\left(\delta_M{}^2\right) + E\left(\delta_X{}^2\right)E\left(\xi_M{}^2\right)$$
$$+ E\left(\delta_X{}^2\right)E\left(\delta_M{}^2\right)$$
$$= Var(\xi_{XM}) + Var(\xi_X)Var(\delta_M) + Var(\xi_M)Var(\delta_X) + Var(\delta_X)Var(\delta_M).$$

(A7)

APPENDIX B

Mplus SYNTAX OF THE PATH MODEL, THE LATENT MODERATED STRUCTURAL EQUATIONS (LMS) METHOD, AND THE RELIABILITY ADJUSTED PRODUCT INDICATOR (RAPI) METHOD

B1: Path Model

```
TITLE:
        Estimate interaction effect with the path model
DATA:
        File=exrep1996.dat;
VARIABLE:
        Names = y xc mc;
        Usevariables=y xc mc xm;
!xc and mc are the mean-centered composites;
!The creation of xc and mc should be conducted outside the Mplus program;

DEFINE:
!xm is the product term of xc and mc;
!grand mean center strategy apply to xm;
        xm=xc*mc;
        center xm (grandmean);
ANALYSIS:
MODEL:
        y ON xc mc xm;
OUTPUT:
        STDYX;
```

B2: Latent Moderated Structural Equations (LMS) Method

```
TITLE:
        Estimate interaction effect with the
        latent moderated structural equations (LMS) method
DATA:
        File=exrep1996.dat;
VARIABLE:
        Names = y xc mc;
        Usevariables=y xc mc ;
!xc and mc are the mean-centered composites;
```

```
!The creation of xc and mc should be conducted outside the Mplus program;

ANALYSIS:
        Type=Random;
        Algorithm=integration;

MODEL:
        fx BY xc;
        fm BY mc;

!Mplus default function for LMS method;
        fxm | fx xwith fm;

        y ON fx fm fxm;

!give labels for latent factor variance;
        fx (vxc);
        fm (vmc);

!give labels for error variance;
        xc (v_exc);
        mc (v_emc);

!specify mean structure
        [fx@0 fm@0];
        [xc@0 mc@0];

   Model Constraint:
! define v_ox and v_om to be the sum of the latent factor variance and error variance,
or the total variance;
        new (v_ox v_om);
        v_ox = vxc + v_exc;
        v_om = vmc + v_emc;

!define the error variance to be the function of reliability and total variance
!in this example, reliability is assumed to be .7;
        v_exc = v_ox*(1-.7);
        v_emc = v_om*(1-.7);

OUTPUT:
        STDYX;
```

B3: Reliability-Adjusted Product Indicator (RAPI) Method

```
TITLE:
        Estimate interaction effect with the
        reliability-adjusted product indicator (RAPI) method
DATA:
        File=exrep1996.dat;
VARIABLE:
        Names = y xc mc;
        Usevariables=y xc mc xm;
!xc and mc are the mean-centered composites;
!The creation of xc and mc should be conducted outside the Mplus program;


DEFINE:
!xm is the product term of xc and mc;
!grand mean center strategy apply to xm;
        xm=xc*mc;
        center xm (grandmean);

MODEL:
!specify the model as shown in Figure 3;
        fx BY xc;
        fm BY mc;
        fxm BY xm;
        y ON fx fm fxm;

!give labels for latent factor variance;
        fx (vxc);
        fm (vmc);
        fxm (vxm);

!give labels for error variance;
        xc (v_exc);
        mc (v_emc);
        xm (v_exm);

!specify mean structure
        [fx@0 fm@0 fxm@0];
        [xc@0 mc@0 xm@0];

   Model Constraint:
! define v_ox, v_om, and v_oxm to be the sum of the latent factor variance and error
variance, or the total variance;
```

```
        new (v_ox v_om v_oxm);
        v_ox = vxc + v_exc;
        v_om = vmc + v_emc;
        v_oxm = vxm + v_exm;

!define the error variance to be the function of reliability and total variance
!in this example, reliability is assumed to be .7;
        v_exc = v_ox*(1-.7);
        v_emc = v_om*(1-.7);
        v_exm = v_ox*.7*v_om*(1-.7)+ v_om*.7*v_ox*(1-.7)
                +v_ox*(1-.7)*v_om*(1-.7);

OUTPUT:
        STDYX;
```