

1

Introduction Draft

2

Jimmy

3

Introduction Draft

Increasing numbers of recent social science research rarely focus on bare bivariate relations among variables; instead, more sophisticated theories with assumptions of intricate effects such as nonlinear and moderation effects are studied as the real world is rarely simple and straightforward (Carte & Russell, 2003; MacKinnon & Luecken, 2008; Cunningham & Ahn, 2019). Numerous research conclude that exercising may help people lose weight, but people may be interested in how, when, for whom, and under what conditions that exercising can do for losing weight. Moderation (or interaction) research can answer such questions by investigating how a third variable (or a groups of additional variables) modifies relations among variables of interest.

In regression models, moderation effects are represented by an interaction term created by multiplying two predictors:

$$Y = b_0 + b_1X + b_2Z + b_3XZ + \epsilon, \quad (1)$$

where b_0 is the intercept, b_1 and b_2 are coefficients for two predictors X and Z , b_3 is the coefficient for the interaction term XZ , and ϵ is the residual term. This general regression model with an interaction term is widely used in most moderation research on observed variables, which means that X and Z are variables that can be directly measured with tangible tools, such as height and weight. Since classical regression model assumes that variables are error-free (i.e., not contaminated by measurement error), parameter estimates (especially the interaction effect) are usually biased when variables are measured with error, which is not uncommon in empirical research. To overcome such problem, latent variables or constructs (e.g., depression), traditionally inferred and measured by a set of observed indicators, are used to control and accommodate measurement error in these observed indicators (Bollen, 2002). Moderation models based on the latent variable modeling (LVM)

framework have been developed to provide reliably true relations among latent constructs (Mueller, 1997; Steinmetz et al., 2011; Cham et al., 2012; Maslowsky et al., 2015).

In this study, a new approach was proposed to estimate latent interaction effects and was developed based on the Two-Stage Path Analysis (2S-PA) method (Lai & Hsiao, 2021): The 2S-PA with Interaction. We carefully reviewed two widely used latent interaction models: Unconstrained Product Indicator (UPI; Marsh et al., 2004) and Reliability-Adjusted Product Indicator (RAPI; Hsiao et al., 2018), and conducted a Monte Carlo simulation study to compare performance of three methods.

The UPI approach was developed to remove complicated constraints in model specification since they increase the probability of erroneous specification and convergence issue (Marsh et al., 2004). Despite easier model implementation, the UPI approach was questioned on arbitrariness in choosing which product indicators (PI) to include as indicators of the latent interaction variables and debatable results of selecting the best strategy under various data distributions. Given increased possible configurations of forming PIs as the number of indicators increase, Marsh et al. (2004) recommended to use a matched-pair strategy in which the first indicator of the first latent variable was multiplied by the first indicator of the second latent variable, and the second indicators of two latent variables were multiplied, and so on for the rest of indicators, and all matched and multiplied PIs indicated to the latent interaction variable. They demonstrated that the matched-pair strategy had better performance under the violation of normal distribution, less convergence issue, and easier model implementation. However, Foldness & Hadtvet (2014) argued that possible variations of forming PIs for the latent interaction variable might bring ambiguity as they lead to different parameter estimations and model fits. In their simulation study, the all-pair UPI (i.e., using all possible pairs of forming PIs) method obtained the best statistical properties and preferable performance with severely non-normal data while reducing the ambiguity of choosing the best configuration. Thus, the lack of consensus on the best way of

implementing UPI resulted in unresolved and controversial issues.

The RAPI method similarly adopts the strategy of forming PI but uses composite scores of raw indicators (i.e., sum or mean score). Specifically, RAPI uses composite scores of mean-centered indicators as the single indicators of latent variables, and the interaction term is constructed by multiplying the single indicators of latent predictors. This method adjusts for the unreliability of composite scores by constraining error variances of observed indicators based on their reliability estimates, thus ensuring that the estimates of the interaction effects are not biased due to measurement error. However, despite its acceptable model complexity and approachable implementation, RAPI might lead to non-positive definite matrices due to negative error variance and inflated interaction effect estimates, under conditions of low reliability ($r = 0.70$) and small sample size ($N = 100$). This suggests that RAPI's performance could be unstable under such conditions.

To account for potential issues in UPI and RAPI, our 2S-PA with Interaction (2S-PA-Int) approach highlights its effectiveness in handling measurement error with small sample sizes while keeping easiness of model complexity and minimal ambiguity of model specification. The 2S-PA-Int involves a three-step process: First, factor scores and their measurement error for each latent variable are first obtained using appropriate psychometric analyses; Second, these factor scores are then used as single indicators of latent predictors in a path analysis framework, with definition variables applied to account for measurement error; Third, the latent interaction term is indicated by multiplying the single indicators of latent predictors and its measurement error is calculated using the derived formula. Lai & Hsiao (2021) had shown that the 2S-PA stood out for its ability to deliver robust, accurate estimates while controlling for measurement error across a range of conditions in their simulation study. Its superiority in convergence rates, bias reduction, and Type I error control, especially in small samples and low reliability scenarios, makes it an advantageous alternative to traditional SEM methods, particularly when dealing with the complexities of

psychological and social science data. Hence the 2S-PA-Int method should inherit the advantages and demonstrate more desirable performance in latent interaction estimation.

To proceed, we first introduced the classical model of latent interaction and then present UPI, RAPI, and 2S-PA with technical details.

A Classical Model of Latent Interaction

Kenny and Judd (1984) first proposed a classical structural model that provided a seminal idea of estimating latent interaction effects. Suppose a simplest case in which two latent predictors and one latent interaction term exist:

$$y = \alpha + \gamma_x \xi_x + \gamma_m \xi_m + \gamma_{xm} \xi_x \xi_m + \zeta, \quad (2)$$

where ξ_x and ξ_m are two latent predictors and their product $\xi_x \xi_m$ forms the interaction term, and α represents the intercept. The ζ term is the disturbance of regression model with a distribution of $\zeta \sim N(0, \psi)$, which captures unobserved factors influencing the dependent variable. The ψ is a scalar representing the variance of ζ . The coefficients γ_x and γ_m represent the main effects, while γ_{xm} represents the interaction effect. The endogenous variable y can be either observed or latent.

The measurement models for ξ (s) in which observed variables are not in mean-deviated forms follow a traditional confirmatory factor analysis (CFA) model:

$$\mathbf{x}_i = \boldsymbol{\tau}_{x_i} + \boldsymbol{\lambda}_{x_i} \xi_x + \boldsymbol{\delta}_{x_i}, \quad (3)$$

For indicators $i = 1, 2, \dots, n$, suppose that a few indicators with a total number n are related to a latent predictor ξ_x . The \mathbf{x}_i is a $q \times 1$ vector of observed exogenous indicators; the ξ_x is a 1×1 scalar representing the latent variable; the $\boldsymbol{\tau}_i$ is a $q \times 1$ vector of intercept terms; $\boldsymbol{\lambda}_{x_i}$ is a $q \times 1$ vector of factor loadings, and $\boldsymbol{\delta}_i$ is a $q \times 1$ vector of measurement errors at the

indicator-level. The measurement error for each indicator, δ_i , follows a normal distribution with a mean of 0 and a variance of θ_i . With an assumption of local independence (i.e., uncorrelated indicators indicating a latent variable), the variance-covariance structure of measurement error of all indicators is a diagonal matrix represented by Θ_δ , where $\Theta_\delta = \text{diag}(\theta_1, \theta_2 \dots \theta_n)$. The measurement model and relevant parameters are similar for ξ_m .

The original Kenny and Judd's model did not include the intercept term α . Jöreskog & Yang (1996) identified this as incorrect and they modified the model with a few assumptions. Three assumptions of the latent interaction model above are commonly imposed with regard to multivariate normal distribution and independence: (1) The measurement errors of indicators, the latent predictor, and the disturbance term in the structural model are multivariate normal, uncorrelated, and independent to each other (i.e., $\phi[\delta, \xi] = 0$; $\phi[\zeta, \xi] = 0$; $\phi[\delta, \zeta] = 0$); (2) All measurement errors are mutually independent and uncorrelated to each other (i.e., $\phi[\delta_1, \delta_2] = 0$; $\phi[\delta_1, \delta_3] = 0$; $\phi[\delta_2, \delta_3] = 0$); (3) The covariance between the two latent predictors (i.e., $\phi[\xi_x, \xi_m]$) is assumed to be freely estimated (not 0) because ξ_{xm} may have a non-normal distribution even though ξ_x and ξ_m are normally distributed (Jöreskog & Yang, 1996).

Algina and Moulder (2001) advanced Jöreskog and Yang's (1996) latent interaction modeling approach by employing a mean-centering strategy. They used the mean-centered indicators to form PIs to indicate the latent interaction term, which offers advantages in interpreting parameter estimates, facilitating model convergence, and reducing bias when estimating interaction effects (Algina & Moulder, 2001; Moulder & Algina, 2002; Marsh et al., 2004). Mean-centering constituent variables within interaction terms diminishes multicollinearity, clarifying the distinct contributions of each variable and their combined influence.

Unconstrained Product Indicator (UPI) [Need a diagram]

Based on Algina and Moulder's model (2001), Marsh et al. (2004) first explored results of removing complicated nonlinear equality constraints required by the classical model, and proposed the influential UPI method for estimating latent interaction effects. For a demonstrative example, suppose ξ_x and ξ_m are related to three indicators respectively:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \tau_{x_1} \\ \tau_{x_2} \\ \tau_{x_3} \end{bmatrix} + \begin{bmatrix} \lambda_{x_1} \\ \lambda_{x_2} \\ \lambda_{x_3} \end{bmatrix} \begin{bmatrix} \xi_x \end{bmatrix} + \begin{bmatrix} \delta_{x_1} \\ \delta_{x_2} \\ \delta_{x_3} \end{bmatrix}, \quad \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} \tau_{m_1} \\ \tau_{m_2} \\ \tau_{m_3} \end{bmatrix} + \begin{bmatrix} \lambda_{m_1} \\ \lambda_{m_2} \\ \lambda_{m_3} \end{bmatrix} \begin{bmatrix} \xi_m \end{bmatrix} + \begin{bmatrix} \delta_{m_1} \\ \delta_{m_2} \\ \delta_{m_3} \end{bmatrix}. \quad (4)$$

Marsh et al. (2004) proposed two strategies of specifying the UPI model: the all-pair UPI and the matched-pair UPI. For the all-pair UPI, the latent interaction term was indicated by all possible configurations of pairs formed by the indicators of ξ_x and ξ_m :

$$\begin{bmatrix} (x_1 - \tau_{x_1})(m_1 - \tau_{m_1}) \\ (x_1 - \tau_{x_1})(m_2 - \tau_{m_2}) \\ (x_1 - \tau_{x_1})(m_3 - \tau_{m_3}) \\ (x_2 - \tau_{x_2})(m_1 - \tau_{m_1}) \\ \dots \\ (x_3 - \tau_{x_3})(m_3 - \tau_{m_3}) \end{bmatrix} = \begin{bmatrix} \lambda_{xm_1} \\ \lambda_{xm_2} \\ \lambda_{xm_3} \\ \lambda_{xm_4} \\ \dots \\ \lambda_{xm_9} \end{bmatrix} \begin{bmatrix} \xi_x \xi_m \end{bmatrix} + \begin{bmatrix} \delta_{xm_1} \\ \delta_{xm_2} \\ \delta_{xm_3} \\ \delta_{xm_4} \\ \dots \\ \delta_{xm_9} \end{bmatrix} \quad (5)$$

where each indicator was the product of two respective indicators of ξ_x and ξ_m , and λ_{xm_i} and δ_{xm_i} are free estimates of factor loadings and measurement errors. The number of product indicators are the product of total numbers of indicators of independent latent predictors. In this example, three observed items (i.e., x_1 to x_3 ; m_1 to m_3) indicating two latent predictors (i.e., ξ_x and ξ_m) together form nine configurations. As for the matched-pair UPI method, the indicators are matched to create the product indicators:

$$\begin{bmatrix} (x_1 - \tau_{x_1})(m_1 - \tau_{m_1}) \\ (x_2 - \tau_{x_2})(m_2 - \tau_{m_2}) \\ (x_3 - \tau_{x_3})(m_3 - \tau_{m_3}) \end{bmatrix} = \begin{bmatrix} \lambda_{xm_1} \\ \lambda_{xm_2} \\ \lambda_{xm_3} \end{bmatrix} \begin{bmatrix} \xi_x \xi_m \end{bmatrix} + \begin{bmatrix} \delta_{xm_1} \\ \delta_{xm_2} \\ \delta_{xm_3} \end{bmatrix} \quad (6)$$

wherein the number of product indicators is greatly reduced due to limited configurations. Marsh et al. (2004) suggested that the matched-pair UPI was more favorable due to two criteria: (1) It used all available information by utilizing every indicators for the latent predictors; (2) It avoided to reuse information by ensuring each indicator was used only once in forming PIs. They stated that the matched-pair UPI ensured the comprehensiveness while maintaining the simplicity of model, and also showed that this method had satisfactory performance in terms of less bias and more robustness under violations of normality assumptions.

Given that the mean of $\xi_x \xi_m$ is not equal to 0 even though ξ_x and ξ_m are assumed to have means of 0 respectively, Marsh et al. (2004) included a mean structure in their UPI model:

$$\kappa = \begin{pmatrix} 0 \\ 0 \\ \phi_{mx} \end{pmatrix} \quad (7)$$

where κ is the mean structure in which the means of ξ_x and ξ_m are assumed to be 0 and ϕ_{mx} as the mean of the latent interaction term is assumed to be freely estimated (i.e., a model parameter).

Recently a more refined method, Double Mean Centering (DMC) strategy, was proposed and shown to eliminate the necessity of including a mean structure, simplify the procedure of model specification and estimation, and demonstrate outstanding performance

under violation of normality assumption (Lin et al., 2010). This method begins with centering all indicators of the latent interaction term, and continues centering the latent interaction variables. Schoemann & Jorgensen (2021) applied DMC to their demonstrative example of latent interaction method by stating that DMC always measures the latent product term and can account for skewed and heterogeneous indicators. We implemented UPI with DMC in this study for efficient model fitting.

The UPI method, compared to the Constrained Product Indicator (CPI) method, is more robust to the violation of normality assumptions of latent predictors with lack of bias and reduced Type I error rate. However, the arbitrariness-complexity dilemma is not well resolved. Although the all-pair UPI perfectly avoided the problem of arbitrary selection of indicators to form PIs, the model complexity will be a serious challenge especially when the number of indicators is large. For example, it is not uncommon that some complicated psychological constructs are indicated by more than 10 items for enough adequacy and depth, and the specified latent interaction term may be indicated by hundreds of items. More items could improve the representation of latent constructs and theoretically increase statistical power for detecting interactions, but also lead to a cumbersome model that negatively impacts interpretability, escalates computational demands, and overfits the sample. The matched-pair UPI can greatly reduce the model complexity, but lead to the problem of selecting indicators for PIs. Marsh et al. (2004) demonstrated the matched-pair strategy with equal-number indicators for the latent predictors, therefore much simplifying the case because each indicator must paired up. As for the case of unequal indicators, which is very likely a common situation in practice for substantive research, multiple observed indicators can be aggregated or grouped into a smaller number of parcels (Jackman et al., 2011), or items with higher reliability are recommended to be maintained to form PIs (Wu et al., 2013). Nevertheless, the great arbitrariness of many alternative strategies results in uncertainty of choosing the best strategy.

Reliability Adjusted Product Indicator (RAPI) [Need a diagram]

Demonstrated in Hsiao et al. (2018), The RAPI method uses a unique strategy of gathering and utilizing information of indicators by using composite scores (e.g., mean score, sum score, etc.) as single indicators of latent predictors and interaction terms instead of multiple PIs, no matter how many indicators latent predictors have.

$$\begin{bmatrix} (x_{comp} - \tau_{x_{comp}}) \\ (m_{comp} - \tau_{m_{comp}}) \\ (x_{comp} \cdot m_{comp} - \tau_{x_{comp} \cdot m_{comp}}) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} \xi_x & \xi_m & \xi_x \xi_m \end{bmatrix} + \begin{bmatrix} \delta_{x_{comp}} \\ \delta_{m_{comp}} \\ \delta_{x_{comp} \cdot m_{comp}} \end{bmatrix}, \quad (8)$$

where x_{comp} and m_{comp} are observed composite scores formed by raw indicators (e.g., $x_{comp} = \Sigma(x_1, x_2, \dots, x_n)$ for $i = 1, 2, \dots, n$). The composite scores indicate their latent variables as single indicators with factor loadings fixed at **1**. Similar to UPI, the measurement errors are represented by δ . The latent interaction term, $\xi_x \xi_m$, is indicated by the product of the composite scores (i.e., $x_{comp} \cdot m_{comp}$). Note that Hsiao et al. (2018) applied the DMC strategy to RAPI, which means that the indicators of latent predictors were first mean-centered (i.e., $x_{comp} - \tau_{x_{comp}}$ and $m_{comp} - \tau_{m_{comp}}$) and the formed single PI of interaction was further mean-centered (i.e., $x_{comp} \cdot m_{comp} - \tau_{x_{comp} \cdot m_{comp}}$).

An important feature of the RAPI method is that it can account for measurement errors of observed indicators when the raw items that produce the composite scores and their products are unreliable, by placing error-variance constraints in the model. The error-variance constraints are computed from the reliability estimates of the raw items. Given that all the factor loadings are set to 1, the reliability index of the raw items, take $x_1 \dots x_i$ as an example, can be simplified to:

$$\rho_{xx'} = \frac{\hat{\sigma}_{\xi_x}^2}{\hat{\sigma}_{\xi_x}^2 + \hat{\sigma}_{\delta_x}^2}, \quad (9)$$

in which $\rho_{XX'}$ is the reliability index, $\hat{\sigma}_{\xi_x}^2$ represent the sample-estimated variance of latent predictors, and $\hat{\sigma}_{\delta_x}^2$ represents the sample-estimated error variance. The reliability estimates can be obtained by composite reliability formula or other widely used estimates (e.g., Cronbach's α). With simple algebra transformation, the error variance of latent independent predictors ξ_x and ξ_m can be shown as a function of the reliability index (Bollen, 1989):

$$\hat{\sigma}_{\delta_x}^2 = (1 - \rho_{xx'})\hat{\sigma}_x^2, \quad (10)$$

$$\hat{\sigma}_{\delta_x} = \frac{1 - \rho_{xx'}}{\rho_{xx'}}\hat{\sigma}_{\xi_x}^2 \quad (11)$$

in which $\hat{\sigma}_x^2 = \hat{\sigma}_{\xi_x}^2 + \hat{\sigma}_{\delta_x}^2$.

Under the assumption of independently and identically distributed measurement error, the equation for the error-variance constraint of the interaction term $\xi_x\xi_m$ is:

$$\begin{aligned} \hat{\sigma}_{\delta_{xm}}^2 = & \rho_{xx'}\hat{\sigma}_x^2(1 - \rho_{mm'})\hat{\sigma}_m^2 + \\ & \rho_{mm'}\hat{\sigma}_m^2(1 - \rho_{xx'})\hat{\sigma}_x^2 + \\ & (1 - \rho_{xx'})\hat{\sigma}_x^2(1 - \rho_{mm'})\hat{\sigma}_m^2, \end{aligned} \quad (12)$$

and the mathematical derivation and technical details are available in Appendix A of Hsiao et al. (2018).

The RAPI method obviously does not have the limitation of arbitrariness in selecting indicators for the interaction term because there is only one way to specify the PI using composite scores. Besides, the model specification for RAPI is simpler because the final number of total PIs is always equal to the total number of latent predictors plus the

interaction terms. By accounting for measurement error, RAPI is expected to produce less biased estimates of interaction effects and enhanced statistical power. However, RAPI depends heavily on accuracy of estimates of reliability measures. Results may be biased if the reliability coefficients as the error constraints are incorrectly estimated. Besides, the RAPI is developed upon the principles of classical test theory (e.g., Cronbach's α for internal consistency). These assumptions include those about. If assumptions of tau-equivalence or parallel measurement are severely violated, the adjustment of constraining errors by RAPI could be compromised. Hsiao et al. (2021) explored the performance of RAPI on congeneric items with four different reliability estimates in a simulation study, and the RAPI showed the best performance with Cronbach's α . As the reliability estimates were adjusted lower, the RAPI method is more likely to overestimate the population interaction effect and inflate standard error estimates.

Two-stage Path Analysis with Interaction (2SPA) [Need a diagram]

The 2SPA was first proposed by Lai and Hsiao (2021), which similarly uses reliability-adjustment technique to account for measurement error in multiple congeneric items. One major improvement of the 2SPA method is that it allows the estimated reliability to be observation-specific and therefore can apply to ordered categorical items account for non-normally distributed (Lai & Hsiao 2021; Lai et al., 2023). Besides, given that a full SEM that estimates measurement and structural models simultaneously generally requires a certain magnitude of sample size and reasonable number of model parameters to ensure convergence rate. The 2S-PA simplifies the model by separating steps of specifying and estimating the measurement and structural model, therefore alleviating computational burden and improving the stability of parameter estimation. We focused on continuous indicators in this study.

At first stage of 2SPA, researchers obtain factor scores (i.e., \hat{F}) of each observation p for latent predictors, $\hat{\xi}_{x_p}$ and $\hat{\xi}_{m_p}$, and individual-specific reliability, $\hat{\rho}_p$, for $p = 1, 2, \dots, k$.

Similar to RAPI, the factor scores of latent predictors form a PI for the interaction term

$\hat{\xi}_{x_p} \hat{\xi}_{m_p}$:

$$\begin{bmatrix} \hat{F}_{x_{ip}} \\ \hat{F}_{m_{ip}} \\ \hat{F}_{xm_{ip}} \end{bmatrix} = \begin{bmatrix} \tau_{x_{ip}} \\ \tau_{m_{ip}} \\ \tau_{xm_{ip}} \end{bmatrix} + \begin{bmatrix} \lambda_{x_{ip}} \\ \lambda_{m_{ip}} \\ \lambda_{xm_{ip}} \end{bmatrix} \begin{bmatrix} \xi_{x_p} & \xi_{m_p} & \xi_{x_p} \xi_{m_p} \end{bmatrix} + \begin{bmatrix} \delta_{x_{ip}} \\ \delta_{m_{ip}} \\ \delta_{xm_{ip}} \end{bmatrix}, \quad (13)$$

in which the factor scores \hat{F}_{x_p} , \hat{F}_{m_p} and $\hat{F}_{xm_{ip}}$ are single indicators of the respective latent variables. Note that the intercepts, factor loadings, and error variances are all model parameters to be freely estimated.

Now suppose that raw indicators of a latent predictor ξ are continuous variables with normal distributions, and let $\hat{\sigma}_F$ be the corresponding standard error of that factor score. Then the error variance is $\hat{\sigma}_F^2$ and it is a constant for all observations under the assumption of continuous indicators with normal distributions. Therefore the symbols just discussed do not keep the observation-specific subscript. Now the error-variance constraints can be computed as $(1 - \rho)\hat{\sigma}_F^2$ in which $\hat{\sigma}_F^2$ is the constant variance of factor score for all observations. Similar constraints are applied to the other latent predictors. The error variance constraint for the interaction term, $\hat{\sigma}_{F_{xm}}^2$, is defined in the same way as RAPI, such as a linear combination of variance and disturbance of two latent variables and reliability. Alternatively speaking, the RAPI method is a special case of 2SPA where the composite scores are used for continuous items (Lai & Hsiao, 2021). As for categorical items with non-normality, the reliability becomes $\rho'_p = 1 - \hat{\sigma}_{F_p}^2$ where the ρ'_p is individual-specific reliability and the $\hat{\sigma}_{F_p}$ is the estimated standard error of the factor score for the p th individual. In this case, the factor loading needs to be fixed individually to the corresponding reliability ρ'_p , and the error variance should be set to $1 - \hat{\sigma}_{F_p}^2 \cdot \rho'_p$.

We examined performance of three methods on estimating the latent interaction

effects across varied conditions under manipulated sample sizes (100, 250 and 500), factor loading (0.75, 0.80, 0.85), correlations between variables (0, 0.3, and 0.6), and reliability levels (.70, .80, .90). Results showed that the 2S-PA-Int method provided less biased estimates, higher coverage rate, lower RMSE and relative SE bias, and higher convergence rate, especially under the condition of low sample size and low reliability level.