

Croon's Bias-corrected Estimation for Multilevel Structural Equation Models with Latent Interactions

 Kyle Cox^a , Benjamin Kelcey^b and Fangxing Bai^b

^aUniversity of North Carolina at Charlotte; ^bUniversity of Cincinnati

 Moderation effects have long been an important consideration for researchers across an array of substantive areas (e.g., Aiken et al., 1991). In the social sciences and across many other disciplines, theories under investigation often involve latent variables and moderation effects operationalized as interactions between latent variables. Structural equation modeling (SEM) is an effective approach in these settings and when hierarchical or clustered data are present, multilevel SEM (MLSEM) is applicable (Preacher et al., 2016). SEM accounts for the uncertainty associated with latent variables and their interactions thereby avoiding biased parameter estimates and inaccuracies in the related standard errors when measurement error is disregarded (e.g., Asparouhov & Muthén, 2021). These advantageous features hold true for MLSEM in addition to its ability to examine between-, within-, and cross-level relationships (Preacher et al., 2016).

Literature has developed several approaches to analyze MLSEM with latent interactions. Contemporary approaches include the latent moderated structural equation method that directly estimates the latent interaction utilizing maximum likelihood estimation with the expectation maximization algorithm (Klein & Moosbrugger, 2000). This approach simultaneously estimates measurement and structural model parameters which has been shown to produce bias and convergence issues with complex models (e.g., >3 latent variables) or relatively small sample sizes (e.g., S. Smid & Rosseel, 2020). The long-established but product-indicator approach utilizes the products of the indicators of the latent focal predictor and the indicators of the latent moderator to form indicators of the interaction (Kenny & Judd, 1984). While this approach is easily implemented in current software (Schoemann & Jorgensen, 2021) it can quickly become unwieldy as the number of observed variables and measurement model parameters increase. A more recent development has been the extension of Bayesian estimation to single- and multilevel models with latent interactions (e.g., Asparouhov & Muthén, 2021). Bayesian approaches also simultaneously consider the measurement and structural models but tend to be more robust to convergence issues, much more computationally efficient, and typically less biased than latent moderated structural equation approaches

(Asparouhov & Muthén, 2021; Cox & Kelcey, 2021; M. Smid et al., 2020). Prior research has, however, suggested that Bayesian estimation generally requires large samples or informative priors for these advantages to fully materialize in complex models such as those including latent interactions (Cox & Kelcey, 2021).

Factor score regression or path analysis represents an alternative to the simultaneous estimation of structural and measurement models employed by latent moderated structural equation methods and Bayesian approaches. Factor score regression takes a sequential approach to estimation, first estimating the measurement models, then using predicted factor scores to estimate the structural or path models. It is relatively easy to implement, fairly robust to model misspecifications, and performs well with smaller sample sizes and more complex SEMs (Ng & Chan, 2020; Zitzmann & Helm, 2021). Ng and Chan (2020) extended a factor score approach to single-level structural equation models with latent interactions and found improved performance. Factor score approaches have since been extended to MLSEMs with latent interactions (e.g., Zitzmann & Helm, 2021). However, typical factor score Regression disregards the measurement error associated with latent variables resulting in biased coefficient estimates (e.g., Cox & Kelcey, 2021). Extension and adaptations to factor score approaches have been proposed to accommodate latent variable measurement error and subsequent parameter estimate bias, but these extensions do not fully accommodate multilevel latent interactions (Brandt et al., 2020; Hoshino & Bentler, 2013; Wall & Amemiya, 2003).

To address this gap, we derived extensions for Croon's bias-corrected estimation to MLSEMs with latent interactions. Croon's bias-corrected estimation approach (Croon's) has performed well in terms of bias and convergence rates with limited sample sizes in a variety of single-level SEMs (Devlieger & Rosseel, 2017), sequential mediation models (Kelcey, 2019), and when estimating mediated effects in MLSEMs (Authors Redacted, in press; Kelcey et al., 2021). It has also performed well with single-level SEMs with varying complexity and latent interactions (Cox & Kelcey, 2021). Similar to factor score regression, Croon's takes a sequential approach that separates estimation of the measurement and

structural models. The distinguishing feature of Croon's is a method of moments correction to adjust estimates in the structural model to account for latent variable unreliability (Croon, 2002). Below we detail structural and measurement models for two-level SEMs with latent interactions, Croon's estimation process, and the corrections necessary for estimating two-level SEMs with latent interactions. Following these description, we investigate the absolute and relative accuracy of the new Croon's based corrections through several simulation studies.

1. Two-level Moderation Effects

We follow the MLSEM latent moderation effect conceptualizations presented in Preacher et al. (2016) and utilized in Asparouhov and Muthén (2021) such that a two-level model with latent moderation can include interactions within levels, between levels, and across levels. We consider seven two-way interaction hypotheses under the four designs possible for a two-level moderation analysis [$1 \times (1-1)$, $2 \times (1-1)$, $2 \times (2-1)$ and $1 \times (2-1)$] where the numbers in the design notation represent the levels of the moderator, focal predictor variable, and outcome (Preacher et al., 2016).

1.1. Design, Model, and Moderation Effect

A $1 \times (1-1)$ design indicates a level-one moderator (Z_{ij}) with level-one focal variable (X_{ij}) and outcome (Y_{ij}) such that $Z_{ij} \times (X_{ij} - Y_{ij})$. To consider all possible forms of moderation in this design, it is necessary to decompose the focal variable (X_{ij}) and moderator (Z_{ij}) into within and between components such that

$$X_{ij} = X_i + X_j \quad (1)$$

$$Z_{ij} = Z_i + Z_j \quad (2)$$

with ij indicating individual i in cluster j and X_j , Z_j cluster specific latent means for each variable (i.e., the between-level component of the variables). There are three feasible moderation effects under the $1 \times (1-1)$ design: First (A1), Z_i moderates the effect of X_i on Y_{ij} , captured using

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i + \varepsilon_{ij} & \varepsilon_{ij} &\sim N(0, \sigma_Y^2) \\ \beta_{0j} &= \gamma_{00} + \gamma_{01} X_j + \gamma_{02} Z_j + u_{0j} & u_{0j} &\sim N(0, \tau_Y^2) \end{aligned} \quad (3)$$

Z , X , and Y variable components retain their meaning from above with the β_3 coefficient capturing the moderation effect and ε_{ij} , u_{0j} the within- and between-level residuals, respectively. This is a within-level interaction effect as both components of the interaction (i.e., Z_i and X_i) are located at the within level. The second moderated effect (A2) possible in a $1 \times (1-1)$ design occurs when Z_j moderates the effect of X_i on Y_{ij} , captured using

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_j + \varepsilon_{ij} & \varepsilon_{ij} &\sim N(0, \sigma_Y^2) \\ \beta_{0j} &= \gamma_{00} + \gamma_{01} X_j + \gamma_{02} Z_j + u_{0j} & u_{0j} &\sim N(0, \tau_Y^2). \end{aligned} \quad (4)$$

The coefficient associated with the moderated effect remains β_3 but the product term now includes the between-level component of the moderator (Z_j) and the within-level component of the focal variable (X_i). With components of

the interaction located at different levels, A2 represents a cross-level moderation effect. The last moderated effect in the $1 \times (1-1)$ design, designated A3, occurs when Z_j moderates the effect of X_j on Y_{ij} , which is modeled using

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_1 X_i + \beta_2 Z_i + \varepsilon_{ij} & \varepsilon_{ij} &\sim N(0, \sigma_Y^2) \\ \beta_{0j} &= \gamma_{00} + \gamma_{01} X_j + \gamma_{02} Z_j + \gamma_{03} X_j Z_j + u_{0j} & u_{0j} &\sim N(0, \tau_Y^2). \end{aligned} \quad (5)$$

With between-level components of both X and Z now making up the interaction term, we have a between-level moderation effect captured with γ_{03} .

In a two-level MLSEM, the moderator may be also be located at level-two (Z_j) while the focal predictor (X_{ij}) remains unchanged. The $2 \times (1-1)$ design has two types of moderated effects. First (B1), Z_j and can moderate the relationship between X_i and Y_{ij} under the model

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_1 X_i + \beta_2 X_i Z_j + \varepsilon_{ij} & \varepsilon_{ij} &\sim N(0, \sigma_Y^2) \\ \beta_{0j} &= \gamma_{00} + \gamma_{01} X_j + \gamma_{02} Z_j + u_{0j} & u_{0j} &\sim N(0, \tau_Y^2) \end{aligned} \quad (6)$$

with the coefficient of interest β_2 capturing a cross-level moderation effect. It is also possible to consider the moderated effect (B2) of Z_j on the relationship between X_j and Y_{ij} using

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_1 X_i + \varepsilon_{ij} & \varepsilon_{ij} &\sim N(0, \sigma_Y^2) \\ \beta_{0j} &= \gamma_{00} + \gamma_{01} X_j + \gamma_{02} Z_j + \gamma_{03} X_j Z_j + u_{0j} & u_{0j} &\sim N(0, \tau_Y^2). \end{aligned} \quad (7)$$

The coefficient associated with the interaction term is now γ_{03} and captures a between-level moderation effect.

Next, a $2 \times (2-1)$ design with the moderator and focal independent variable measured at level-two produces one moderated effect (C) when Z_j influences the relationship between X_j and Y_{ij} modeled by

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \varepsilon_{ij} & \varepsilon_{ij} &\sim N(0, \sigma_Y^2) \\ \beta_{0j} &= \gamma_{00} + \gamma_{01} X_j + \gamma_{02} Z_j + \gamma_{03} X_j Z_j + u_{0j} & u_{0j} &\sim N(0, \tau_Y^2). \end{aligned} \quad (8)$$

The coefficient associated with the interaction is again γ_{03} and captures a between-level moderation effect.

The last two-level design is a $1 \times (2-1)$ design that includes a focal predictor located at level-two but a moderator located at level-one. For this moderation effect (D), Z_j influences the relationship between X_j and Y_{ij}

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_1 Z_j + \varepsilon_{ij} & \varepsilon_{ij} &\sim N(0, \sigma_Y^2) \\ \beta_{0j} &= \gamma_{00} + \gamma_{01} X_j + \gamma_{02} Z_j + \gamma_{03} X_j Z_j + u_{0j} & u_{0j} &\sim N(0, \tau_Y^2). \end{aligned} \quad (9)$$

The interaction term is located at level-two and the between-level moderation effect is again captured by γ_{03} .

The four multilevel moderation designs noted here and the seven associated moderation effects possible under those designs require six different Croon's corrections. For moderation effects, the corrections for Croon's estimation adjust the covariance between latent variables and latent interactions and are dependent on the location of the variables and interaction (i.e., between or within level). For two-level SEMs with latent interactions, three corrections are needed

for moderation effects at the between level, one correction for the moderation effect at the within level, and two corrections for cross-level moderation effects. Specifically, corrections for between-level moderation effects are necessary for the covariance of the between part of a level-one moderator and the between part of a level-one predictor (A3), level-two moderator and level-two predictor (C), and level-two moderator or predictor and between part of a level-one moderator or predictor (B2 and D). Moderation effects at the within level require corrections for the covariance of the within part of a level-one moderator and within part of a level-one predictor (A1). Cross-level moderation effects require corrections for the covariance of the between part of a level-one moderator and within part of a level-one predictor (A2), and level-two moderator and the within part of a level-one predictor (B1).

1.2. Measurement Models

Before detailing Croon's corrections for two-level SEMs with latent interactions, we present factor models for each latent variable. While previous literature has emphasized observed variables or latent variables with simple measurement models in two-level moderation analyses (e.g., Asparouhov & Muthén, 2021; Preacher et al., 2016), we consider a latent outcome, focal predictor, and moderator measured using multiple indicators. The latent variables of interest may be located at the within- or between level. When the latent variable is measured at level-two and therefore limited to the between level (e.g., Z_j in Equation (6)), a single-level common factor model is employed such that

$$Z_j = \mu_{z_j} + \Lambda_{z_j}\eta_{z_j} + \varepsilon_{z_j}. \quad (10)$$

Here, Z_j are the observed indicators for the latent η_{z_j} , Λ_{z_j} represents the factor loadings, μ_{z_j} the indicator intercepts, and ε_{z_j} the error terms. To fix the scale the variance of η_{z_j} can be set to one or the factor loading of the first indicator can be set to one. We utilized the latter method to set the scale and utilized three indicators per latent variable in our models and simulation studies. The factor model for η_{z_j} is fit using maximum likelihood with factor scores for η_{z_j} determined using the regression predictor method and represented by $\tilde{Z}_j$. A similar model is utilized for latent focal predictors measured at level-two (e.g., η_{x_j} for X_j).

For latent variables measured at level-one, it is necessary to employ a multilevel factor model allowing each variable to be decomposed into between- and within-level components. For example, the focal independent variable X_{ij} is decomposed into $\eta_{X_j}^{L2}$ and $\eta_{X_i}^{L1}$ such that

$$X_{ij} = \mu_{X_j} + \Lambda_{X_j}^{L2}\eta_{X_j}^{L2} + \Lambda_{X_i}^{L1}\eta_{X_i}^{L1} + \varepsilon_{X_j}^{L2} + \varepsilon_{X_i}^{L1}. \quad (11)$$

In the multilevel factor model i and j index the individuals and clusters, respectively, X_{ij} represents the indicators for the latent variable, the latent variable has a between-level component represented by $\eta_{X_j}^{L2}$ and a within-level component $\eta_{X_i}^{L1}$. $\Lambda_{X_j}^{L2}$ and $\Lambda_{X_i}^{L1}$ are the corresponding factor loadings, μ_{X_j} are the intercepts for each cluster, and $\varepsilon_{X_j}^{L2}$ and $\varepsilon_{X_i}^{L1}$ are the between- and within-level error terms. Assigning

unit variances to the between- and within-level factors can set the scale but we again set the first factor loading of both $\eta_{X_j}^{L2}$ and $\eta_{X_i}^{L1}$ to set the scale. Similar formulations are appropriate for other latent variables measured at level-one (e.g., the latent outcome Y_{ij} and latent moderator Z_{ij}).

While single-level factor scoring methods (i.e., regression predictor method) are suitable for variables located entirely at level-two (i.e., Z_j in Equation (6)), we must consider the clustering of observed indicators when factor scores are determined for level-one variables that vary within- and between levels (i.e., X_{ij} in Equation (6); Kelcey et al., 2021). For example, the between-level factor scores for the focal predictor (X_{ij}) in Equation (6) can be found with

$$\tilde{X}_{.j}^{L2} = A_{X_1}^{L2}x_{1j}^{L2} + A_{X_2}^{L2}x_{2j}^{L2} + A_{X_3}^{L2}x_{3j}^{L2} \quad (12)$$

The predicted factor scores for cluster j are represented with $\tilde{X}_{.j}^{L2}$, $x_{.j}^{L2}$ are the between-level components of the focal predictor indicators, and $A_{X_j}^{L2}$ are the between-level factor score matrix weights. To predict between-level factor scores for the between-level component of a latent variable (e.g., factor scores for the focal predictor $\tilde{X}_{.j}^{L2}$), the latent between-level components of the within-level indicators (x^{L2}) are required along with the between-level covariance matrix of the focal predictor indicators and between-level factor loadings but these are available using the multilevel factor model (Equation (11)). Put differently, the indicator values at the between level of a multilevel latent variable are not directly observed so they must also be predicted while indicator values for single-level latent variables are observed and thus directly used to predict factor scores. While alternative (e.g., empirical Bayes) and multivariate approaches are available, we use the univariate cluster means approach to estimate between-level components of indicators. This approach simply uses the observed indicator cluster means to estimate between-level components of indicators (Kelcey et al., 2021). Of pertinence to our corrections, we can now incorporate the reliability of between-level factor scores (R^{L2}) in corrections involving between-level components of variables measured at the within level.

2. Croon's Bias-corrected Estimation

Croon's estimation can be classified as a Structural After Measurement approach (Rossee & Loh, in press) as it first estimates measurement models, then structural models with a Croon's based correction (Croon, 2002). Croon's tracks bias introduced into the structural or path model by the unreliability associated with the latent variables and then corrects for that uncertainty using measurement model results (Croon, 2002; Kelcey et al., 2021). The process of Croon's estimation does parallel factor score regression but, importantly, Croon's estimation includes a correction to structural relationships to account for measurement error. More specifically, Croon's bias-corrected estimation employs a four step process: (a) a factor model is fit for each latent variable using, for example maximum likelihood, and then factor scores are predicted for each latent variable; (b) the

variance-covariance matrix is estimated using the aforementioned factor scores; (c) a method of moments type correction is constructed using results from the measurement models and applied to the variance-covariance matrix produced in step (b); and (d) the structural or path model is estimated using the corrected variance-covariance matrix.

The correction of the factor score covariance matrix (step (c) above) is the crucial and distinguishing component of a Croon's estimation. The method of moments correction adjusts the factor score covariance matrices at each level to account for latent variable measurement error (Croon, 2002; Kelcey et al., 2021). Croon's based corrections utilize the reliability of latent variables as derived from the measurement models to adjust the covariances proportional to the magnitude of the measurement error. Because the corrections are applicable to the covariance between two variables they are structured in variable pairs: (a) individual-level covariance between two individual-level variables, (b) cluster-level covariance between two individual-level variables, (c) covariance between a cluster-level variable and individual-level variable, and (d) covariance between two cluster-level variables. The covariance between a cluster-level variable and the between-level component of an individual-level variable is corrected using the cross-level correction (c). For a more detailed explanation of each correction associated with Croon's estimation of MLSEMs see Kelcey et al. (2021) and Devlieger and Rosseel (2020), for single-level latent interactions see Cox and Kelcey (2021), and for single-level mediation see Devlieger and Rosseel (2017) and Kelcey (2019).

Cox and Kelcey (2021) applied the Croon's correction approach to estimating latent interactions in single-level SEMs. The process and logic follow the previously detailed steps but the corrections involving the covariance between a variable and the interaction terms considered the measurement models of all three variables. Croon's has also been successfully applied across other single-level SEMs (e.g., Devlieger et al., 2016; Devlieger & Rosseel, 2017; Kelcey, 2019; Loncke et al., 2018) and various multilevel contexts (Authors Redacted, in press; Devlieger & Rosseel, 2020; Kelcey et al., 2021). Across these settings Croon's produced nearly unbiased coefficient estimates and relative to other prevalent estimators (i.e., ML and FS), Croon's was often more efficient and more likely to converge. The sequential approach of Croon's is well suited for complex models including MLSEMs with latent interactions when sample sizes are limited (Authors Redacted, in press; Kelcey et al., 2021).

The robustness of Croon's to model misspecifications and its performance under various measurement model conditions is an area of ongoing research. Hayes and Usami (2020a, 2020b) investigated multidimensionality and correlated errors under Croon's method, and Devlieger and Rosseel (2020) and Kelcey et al. (2021) investigated the influence of item intraclass correlation on the accuracy and sensitivity of Croon's method in multilevel settings. In terms of Croon's robustness to model misspecifications, Cox and Kelcey (2022) considered the sensitivity of Croon's method

to non-normality in indicators, missing path values in the structural models, indicators loading on the incorrect factor, and missing cross-loadings in the measurement model. Hayes and Usami (2020b) also considered Croon's performance with misspecifications in the structural model. While these works and others (e.g., Rosseel & Loh, in press) have generally found Croon's based approaches to be more robust to model misspecifications than traditional maximum likelihood estimators, Robitzsch (2022) found this to be true only for structural model misspecifications and not for misspecifications involving cross-loadings and residual correlations.

3. Croon Corrections for Multilevel Latent Interactions

Using the measurement models and the structural models detailed previously, we extend Croon's estimation to accommodate latent interactions in two-level SEMs. We detail the six different Croon based corrections needed to accommodate the seven latent moderation effects possible in two-level SEMs. We begin with interactions involving variable components at the between level (moderation effects A3, C, B2, and D captured with γ_{03}) then detail corrections when variable components are located at the within level (moderation effects A1), and conclude with cross-level moderation effects (A2, and B1). With each correction we present simulation study results to illustrate the accuracy, efficiency, and feasibility of Croon's estimation for MLSEMs with different moderation effects.

As with previous extensions of Croon's estimation, our corrections do not accommodate random slopes, assume cross-level invariance on the factor loadings and normally distributed indicator residuals that are uncorrelated across factors. Relaxing these assumptions are ongoing areas of research (e.g., Blinded, 2022) but are beyond the scope of this study. We discuss specific directions for future research in the Discussion section.

3.1. Between Part of Level-one Moderator $\times$ between Part of Level-one Predictor (A3)

Conceptually, Croon's corrections adjust the factor score covariance matrices at each level using the reliability of the latent variables to account for latent variable measurement error (Croon, 2002; Kelcey et al., 2021). For the moderation effect in Equation (5) (A3), we must correct the covariance of the interaction involving the between part of the level-one moderator (Z_j^{L2}) and the between part of the level-one predictor (X_j^{L2}) with the outcome (Y_j^{L2}). More formally, we must correct $cov(\eta_{Z_j}^{L2}, \eta_{X_j}^{L2}, \eta_{Y_j}^{L2})$ where $cov(\hat{\eta}_{Z_j}^{L2}, \hat{\eta}_{X_j}^{L2}, \hat{\eta}_{Y_j}^{L2}) = cov(\tilde{Z}_j^{L2} \cdot \tilde{X}_j^{L2}, \tilde{Y}_j^{L2})$ with $\tilde{Y}_j^{L2}$ the between-level factor scores of the outcome and $\tilde{Z}_j^{L2} \cdot \tilde{X}_j^{L2}$ representing the product of factor scores from $\tilde{Z}_j^{L2}$ and $\tilde{X}_j^{L2}$. If we assume a simple structure (i.e., no cross-loadings and conditional independence of the measurement residuals across factor models), a

corrected estimate of the covariance between the interaction and the outcome can be obtained using

$$\text{cov}(\eta_{Z_j}^{L2} \eta_{X_j}^{L2}, \eta_{Y_j}^{L2}) = \frac{\text{cov}(\hat{\eta}_{Z_j}^{L2} \hat{\eta}_{X_j}^{L2}, \hat{\eta}_{Y_j}^{L2})}{\mathbf{A}_{Z_j}^{L2} \mathbf{R}_{Z_j}^{L2} \mathbf{A}_{X_j}^{L2} \mathbf{R}_{X_j}^{L2} \mathbf{A}_{Y_j}^{L2} \mathbf{R}_{Y_j}^{L2}} \quad (13)$$

This correction parallels the correction for latent interactions in single-level SEMs (Cox & Kelcey, 2021). The estimated covariance between the predicted factor scores ($\text{cov}(\hat{\eta}_{Z_j}^{L2} \hat{\eta}_{X_j}^{L2}, \hat{\eta}_{Y_j}^{L2})$) is divided by the product of the factor score matrix ($\mathbf{A}$) and factor loading matrix ($\mathbf{\Lambda}$) from each latent variable involved in the covariance ($\eta_{Z_j}^{L2}$, $\eta_{X_j}^{L2}$, and $\eta_{Y_j}^{L2}$). Unique to MLSEM interactions is the $\mathbf{R}^{L2}$ term representing the vector of mean indicator reliabilities. The first two components ($\mathbf{A}$, $\mathbf{\Lambda}$) adjust for the unreliability of the factors while $\mathbf{R}^{L2}$ properly adjusts for the unreliability of indicator cluster means.

After correcting covariances it is necessary to address the variances of the within- and between-level components of all the latent variables. Fixing the scale by setting the variances to one negates the need to correct the variance terms. If one sets the scale of latent variable components using the variance of the first indicator, a corrected variance is available from the factor models. For example, we utilized the estimated variance of the within- and between-level components of all the latent variables from the factor models. However, neither of these solutions are applicable to the variance of the latent interaction (e.g., $\text{var}(\eta_{Z_j}^{L2} \eta_{X_j}^{L2})$) as it is not estimated in a measurement model. For latent variable interactions at both the within- and between level, the variance can be estimated using the variance of the product of two random variables (Bohrnstedt & Goldberger, 1969; Cox & Kelcey, 2021) such that

$$\text{var}(\eta_{Z_j}^{L2} \eta_{X_j}^{L2}) = \text{var}(\eta_{Z_j}^{L2}) \text{var}(\eta_{X_j}^{L2}) + \text{cov}(\eta_{Z_j}^{L2}, \eta_{X_j}^{L2})^2. \quad (14)$$

Here, it is assumed that the mean of each latent variable component is centered at zero. Note that the corrected variance of the latent interaction (e.g., $\text{var}(\eta_{Z_j}^{L2} \eta_{X_j}^{L2})$ in Equation (14)) is calculated using corrected variance (e.g., $\text{var}(\eta_{X_j}^{L2})$ and $\text{var}(\eta_{Z_j}^{L2})$) and covariance (e.g., $\text{cov}(\eta_{Z_j}^{L2}, \eta_{X_j}^{L2})$) components.

3.1.1 Simulation

To assess the performance of Croon's estimation with latent interactions in MLSEMs we conducted a series of simulation studies. Data were generated using R (R Development Core Team, 2019) based on the corresponding multilevel moderation model and measurement models. Analyses were completed using Mplus 8.6 (Muthén & Muthén, 1998–2021) and the MplusAutomation package in R (Hallquist & Wiley, 2018). For example, simulated data for examining Croon's estimation for an A3 moderation effect, the between part of the level-one moderator (Z_j^{L2}) and the between part of the level-one predictor (X_j^{L2}), were generated in R using Equation (11) for the three indicators of each latent variable and Equation (5) for structural components (see Figure 1). Maximum likelihood was used in Mplus to estimate the factor models and structural models for both Croon's and

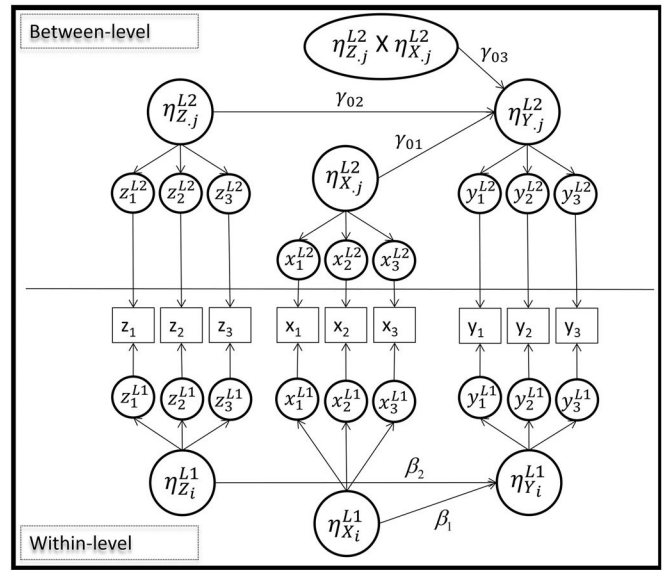


Figure 1. Conceptual model for MLSEM with latent interaction composed of the between part of a level-one moderator and between part of a level-one predictor (A3).

factor score regression approaches. Results were then compiled from Mplus outputs files.

The first factor loading of each indicator was fixed at 1.0 to set the scale of the latent variable. Factor loadings of the second and third indicator at the within level were set at 1.5 and 0.5, respectively. We switched these values at the between level such that the second indicator loading was 0.5 and the third loading was 1.5. Note that the multiple (three) indicator measurement models considered in our analyses are common in methodological and substantive research but also more complex than those used in many previous simulation studies of MLSEMs with latent interactions (e.g., single indicator models; Asparouhov & Muthén, 2021; Preacher et al., 2016). We constrained all indicator residuals to be positive to avoid inflated non-positive definite covariance matrixes from negative indicator residuals. Path coefficients for each simulation were set at $\beta_1 = \gamma_0 = 0.2$. The outcome variance decomposition in the structural models across simulation studies was $\sigma_Y^2 = 0.8$ and $\tau_Y^2 = 0.2$. Error variance of the multilevel factor models had similar decomposition with the variance of $\epsilon_{X_i}^{L1}$ set to 0.8 and $\epsilon_{X_j}^{L2}$ set to 0.2 while error variance of single-level factor models was set to 1.0.

We varied cluster sample size ($n_2 = 90, 60, 30$) and individual per cluster sample size ($n_1 = 50$ and 20) to reflect limited sample size conditions in which prevalent and default estimators for MLSEMs (i.e., maximum likelihood) have demonstrated poor performance especially when latent interactions are present (Devlieger & Rosseel, 2020; Kelcey et al., 2021; Rosseel, 2020; S. Smid & Rosseel, 2020). The fully crossed design produced six sample size conditions for each of the six corrections. Each simulation study utilized 500 data sets with coefficient estimates captured using Croon's estimation and an uncorrected factor score regression approach (FS). Default maximum likelihood estimation

was excluded as a comparison estimator because initial simulation study results even those with larger sample sizes found it failed to converge in the vast majority of conditions, took substantially more time to converge than Croon's or FS, and often yielded extreme parameter estimates when results were produced. Maximum likelihood was simply not a feasible estimator for the types of conditions we considered for the two-level SEMs with latent interactions.

3.1.2. Results

Estimator performance was evaluated using convergence failure rate, bias, and efficiency. These metrics indicate the absolute performance of Croon's estimation while a comparison of these metrics to the uncorrected FS approach helps evaluate the newly developed corrections. Below we detail convergence failure rate, bias, and efficiency then present related results followed by similar presentations for each multilevel moderation effect.

3.1.2.1. Convergence Failure Rate. Convergence failure rate indicates the rate at which an estimator failed to converge (i.e., provided no results). Convergence failure for Croon's estimation and FS typically stemmed from a non-positive definite matrix in the measurement model or, for Croon's only, a variable with a variance approaching zero after the correction. Croon's estimation avoided convergence issues (i.e., reliably provided results) even at the smallest sample sizes (i.e., $n_2 = 30$). It failed to produce results in around 3% of the simulation runs at this sample size and the convergence failure rate fell to near 0 as n_2 increased to 90 or more (see Table 1). Overall, both estimators performed well with the A3 multilevel moderation effect model.

3.1.2.2. Bias. To evaluate the accuracy of Croon's estimation with our new corrections for multilevel latent interactions, we tracked bias in coefficient estimates across all simulated data sets. Bias was calculated as the difference between the estimated coefficient value and the true value within each data set. We found the average absolute bias at the

Table 1. Summary of estimator performance with latent interaction A3 in terms of convergence failure rate, bias, and efficiency across sample size conditions.

Estimator	n_1	n_2	Fail converge	Bias		SD		RMSE	
				L2	L1	L2	L1	L2	L1
Croon	50	30	0.006	0.016	0.002	0.113	0.043	0.114	0.043
		60	0.008	0.006	0.003	0.078	0.031	0.078	0.032
		90	0	0.005	0.002	0.063	0.025	0.064	0.026
	20	30	0.032	0.013	0.006	0.121	0.067	0.121	0.067
		60	0.018	0.011	0.005	0.082	0.047	0.083	0.047
		90	0.004	0.006	0.006	0.066	0.039	0.066	0.04
FS	50	30	0	0.028	0.026	0.098	0.03	0.102	0.04
		60	0	0.025	0.027	0.064	0.022	0.069	0.034
		90	0	0.027	0.027	0.05	0.018	0.057	0.032
	20	30	0.006	0.029	0.028	0.099	0.049	0.103	0.056
		60	0	0.034	0.027	0.065	0.034	0.073	0.044
		90	0.002	0.032	0.028	0.051	0.028	0.06	0.04

Note: Fail converge represents the rate estimators failed to converge. Estimators are abbreviated with Croon for Croon's bias-corrected estimation and FS for the uncorrected factor score regression approach.

within- (L1) and between level (L2) using all coefficient estimates at the given level. This produced two summative values capturing bias for each estimator (i.e., within-level bias and between-level bias)

Croon's estimation performed well in terms of bias with the A3 multilevel moderation model (see Table 1). These results indicate the newly derived corrections involving an A3 multilevel moderation effect are accurate. Croon's estimation demonstrated decreased bias as n_2 increased but this trend was not present in FS estimates. The FS approach fails to consider latent variable unreliability which produces biased coefficient estimates across sample size conditions. We found greater amounts of bias in between-level coefficient estimates when compared to within-level coefficient estimates. This result was expected due to differences in sample size. To understand these results in a broader context, recall that each coefficient value was set at 0.2 indicating relative bias across all models can be found using $bias/0.2$. For example, bias of .03, .02, and .01 is equivalent to relative bias of 0.15, 0.1, and 0.05 indicating Croon's estimation was near or below conventional guidelines for relative bias of <0.05 (Hoogland & Boomsma, 1998) across conditions for A3 latent interactions.

3.1.2.3. Efficiency. Estimator accuracy is vital but efficiency or the precision of estimates is also an important criterion to evaluate estimator feasibility. We used the standard deviation of coefficient estimates across replications to gauge estimator efficiency. Specifically, we averaged the standard deviation of coefficient estimates at each level to measure estimator efficiency (see SD in Table 1). Croon's estimation again performed comparatively well. The FS approach also performed well on this metric as FS approaches typically produce efficient estimates due to their exclusion of latent variable unreliability. Our results reflect this expectation.

To consider bias and efficiency concurrently, we calculated the root mean square error (RMSE) for each coefficient estimate and tracked the average RMSE for each estimator across the between- and within level (see Table 1). Croon's estimation demonstrated a good balance of accuracy and efficiency when estimating the A3 multilevel moderation model. In totality, Croon's estimation demonstrated dependable (i.e., low convergence failure rate), accurate (i.e., small bias), and efficient (i.e., small SD) estimates of the A3 multilevel moderation model. In the remaining corrections for the different types of multilevel moderation effects in two-level SEMs and corresponding simulation study results, we emphasize the distinguishing features of each correction, any changes to simulation conditions, and results that diverge from other findings.

3.2. Level-two Moderator $\times$ Level-two Predictor (C)

For the moderation effect (C) in Equation (8) (see Figure 2a), we must correct the covariance of an interaction involving a level-two moderator (Z_j) and a level-two predictor (X_j) with the outcome (Y_j^{L2}). Following a similar process and assumptions, a corrected estimate of the covariance

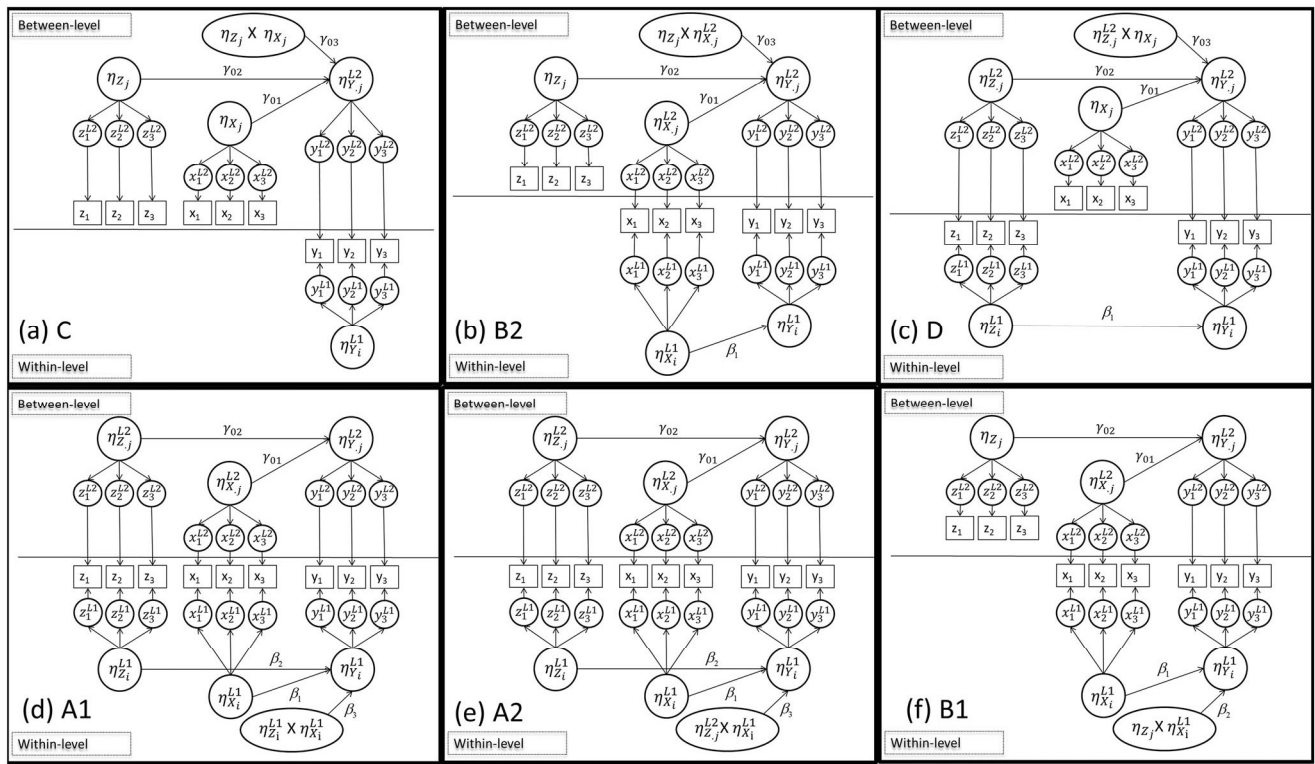


Figure 2. Conceptual models for MLSEMs with latent interactions for various moderation effects (a) C, (b) B2 (c) D, (d) A1, (e) A2, and (f) B1.

between this latent interaction and the latent outcome ($\text{cov}(\eta_{Z_j} \eta_{X_j}, \eta_{Y_j}^{L2})$) can be obtained using

$$\text{cov}(\eta_{Z_j} \eta_{X_j}, \eta_{Y_j}^{L2}) = \frac{\text{cov}(\hat{\eta}_{Z_j} \hat{\eta}_{X_j}, \hat{\eta}_{Y_j}^{L2})}{\mathbf{A}_{Z_j} \mathbf{A}_{X_j} \mathbf{A}_{Z_j} \mathbf{A}_{X_j} \mathbf{A}_{Y_j}^{L2} \mathbf{R}_{Y_j}^{L2} \mathbf{A}_{Y_j}^{L2}} \quad (15)$$

with variance of the interaction found using

$$\text{var}(\eta_{Z_j} \eta_{X_j}) = \text{var}(\eta_{Z_j}) \text{var}(\eta_{X_j}) + \text{cov}(\eta_{Z_j}, \eta_{X_j})^2. \quad (16)$$

Notice the correction is a simplified form of the one utilized for the A3 multilevel moderation effect. This correction does not include indicator reliability terms ($\mathbf{R}^{L2}$) for Z_j or X_j as they are not necessary for variables measured at the second level of the model. The indicators of Z_j and X_j are located at level-two and therefore are directly observed. Variance of the latent interaction is again estimated using the variance of the product of two random variables (Bohrnstedt & Goldberger, 1969; Cox & Kelcey, 2021).

We conducted a similar simulation study to examine Croon's estimation for a MLSEM that included a type C multilevel moderation effect. Croon's estimation again performed well in terms of dependability, accuracy, and efficiency (see Table 2). Of note, the model only includes latent variables at the between level so level-one results are necessarily omitted. As for Croon's, it performed well in terms of convergence failure, bias, and efficiency. Convergence failure rates were a bit higher than those found estimating a MLSEM with an A3 moderated effect but quickly dropped as n_2 exceeded 30 and never exceeded 6%. This result likely stems from the location of all variables (and therefore coefficient estimates) at the between level. For comparison, the FS approach had low convergence failure rates and small SD of

Table 2. Summary of estimator performance with latent interaction C in terms of convergence failure rate, bias, and efficiency across sample size conditions.

Estimator	n_1	n_2	Fail converge	Bias L2	SD L2	RMSE L2
Croon	50	30	0.058	0.015	0.143	0.145
		60	0.006	0.015	0.094	0.095
		90	0	0.012	0.072	0.074
	20	30	0.044	0.024	0.15	0.152
		60	0.016	0.015	0.103	0.104
		90	0	0.007	0.076	0.077
FS	50	30	0.006	0.028	0.126	0.129
		60	0	0.034	0.077	0.085
		90	0	0.031	0.06	0.068
	20	30	0.01	0.04	0.119	0.126
		60	0	0.035	0.078	0.086
		90	0	0.033	0.06	0.069

Notes: Fail converge represents the rate estimators failed to converge. Estimators are abbreviated with Croon for Croon's bias-corrected estimation and FS for the uncorrected factor score regression approach.

coefficient estimates but consistent bias ≈ 0.03 , slightly higher than found with the A3 moderation effects model.

3.3. Level-two Moderator (or Predictor) $\times$ between Part of Level-one Moderator (or Predictor) (B2 or D)

A single correction is applicable to the multilevel moderation effects in Equations (7) and (9) (see Figure 2b and c). In both cases (i.e., moderation effects B2 and D), a latent variable measured at level-two covaries with the between part of a level-one latent variable. For the B2 moderation effect, we have a level-two moderator (Z_j) influencing the relationship of the between-level component of the level-one

Table 3. Summary of estimator performance with latent interactions B2 in terms of convergence failure rate, bias, and efficiency across sample size conditions.

Estimator	n_1	n_2	Fail converge	Bias		SD		RMSE	
				L2	L1	L2	L1	L2	L1
Croon	50	30	0.036	0.015	0.007	0.129	0.04	0.13	0.04
		60	0	0.009	0.003	0.089	0.028	0.09	0.028
		90	0.006	0.005	0.005	0.069	0.022	0.069	0.023
	20	30	0.04	0.019	0.011	0.136	0.06	0.137	0.061
		60	0.016	0.013	0.011	0.088	0.045	0.09	0.046
		90	0.002	0.005	0.011	0.074	0.035	0.075	0.037
FS	50	30	0.004	0.031	0.046	0.107	0.032	0.111	0.056
		60	0	0.029	0.043	0.071	0.022	0.077	0.048
		90	0	0.027	0.044	0.055	0.018	0.062	0.048
	20	30	0.01	0.034	0.044	0.114	0.049	0.119	0.066
		60	0.004	0.036	0.046	0.07	0.037	0.079	0.059
		90	0.002	0.032	0.045	0.058	0.029	0.066	0.054

Notes: Fail converge represents the rate estimators failed to converge. Estimators are abbreviated with Croon for Croon's bias-corrected estimation and FS for the uncorrected factor score regression approach.

focal predictor (X_j) on the outcome (Y_{ij}). For the D moderation effect, the location and type of the moderator and focal predictor switch with the between-level component of the level-one moderator (Z_j) influencing the relationship of a level-two focal predictor (X_j) on the outcome (Y_{ij}). We use the B2 multilevel moderation effect to present the correction for a covariance involving a level-two latent variable and the between part of a level-one latent variable ($\text{cov}(\eta_{Z_j}\eta_{X_j}^{L2}, \eta_Y^{L2})$). The corrected covariance is

$$\text{cov}(\eta_{Z_j}\eta_{X_j}^{L2}, \eta_Y^{L2}) = \frac{\text{cov}(\hat{\eta}_{Z_j}\hat{\eta}_{X_j}^{L2}, \hat{\eta}_Y^{L2})}{\mathbf{A}_{Z_j}\mathbf{A}_{X_j}^{L2}\mathbf{R}_{X_j}^{L2}\mathbf{A}_{Z_j}^{L2}\mathbf{A}_{X_j}^{L2}\mathbf{A}_{Y_j}^{L2}\mathbf{R}_{Y_j}^{L2}\mathbf{A}_{Y_j}^{L2}} \quad (17)$$

With variance of the interaction such that

$$\text{var}(\eta_{Z_j}\eta_{X_j}^{L2}) = \text{var}(\eta_{Z_j})\text{var}(\eta_{X_j}^{L2}) + \text{cov}(\eta_{Z_j}, \eta_{X_j}^{L2})^2 \quad (18)$$

This correction follows the previous corrections with the biased factor score covariance representing the latent interaction and latent outcome ($\text{cov}(\hat{\eta}_{Z_j}\hat{\eta}_{X_j}^{L2}, \hat{\eta}_Y^{L2})$) divided by components that capture unreliability (i.e., $\mathbf{A}$, $\mathbf{A}^{L2}$, and $\mathbf{R}^{L2}$). The $\mathbf{R}^{L2}$ terms are included for the latent variables located at level-one (η_Y and η_X) but are unnecessary for the level-two latent variable (η_Z).

Results for the simulation study examining estimation of the B2 multilevel moderation model can be found in Table 3. Convergence failure rates remained at or below 4% for Croon's estimation even at the smallest sample sizes while increases in sample size continued to reduce bias and increase efficiency. The FS approach continued to demonstrate consistently biased coefficient estimates.

3.4. Within Part of Level-one Moderator $\times$ within Part of Level-one Predictor (A1)

The corrections detailed thus far have involved the covariance of a latent interaction with another latent variable at the between level (see Figures 1 and 2). There are also unique corrections necessary for the moderation effects at the within level and across levels. We begin with a correction for an A1 moderation effect that involves the within part of a level-one moderator (Z_i) and within part of a

level-one predictor (X_i ; see Figure 2d). This moderation effect occurs entirely at the within level. More formally, we must correct $\text{cov}(\eta_{Z_i}^{L1}\eta_{X_i}^{L1}, \eta_{Y_{ij}}^{L1})$ where $\text{cov}(\hat{\eta}_{Z_i}^{L1}\hat{\eta}_{X_i}^{L1}, \hat{\eta}_{Y_{ij}}^{L1}) = \text{cov}(\tilde{Z}_i^{L1} \cdot \tilde{X}_i^{L1}, \tilde{Y}_{ij}^{L1})\text{cov}(\tilde{Z}_i^{L1} \cdot \tilde{X}_i^{L1}, \tilde{Y}_{ij}^{L1})$ with $\tilde{Z}_i^{L1} \cdot \tilde{X}_i^{L1}$ representing the product of factor scores from $\tilde{Z}_i^{L1}$ and $\tilde{X}_i^{L1}$ and $\tilde{Y}_{ij}^{L1}$ representing the within-level factor scores of the outcome. Retaining our process and assumptions, a corrected covariance can be obtained using

$$\text{cov}(\eta_{Z_i}^{L1}\eta_{X_i}^{L1}, \eta_{Y_{ij}}^{L1}) = \frac{\text{cov}(\hat{\eta}_{Z_i}^{L1}\hat{\eta}_{X_i}^{L1}, \hat{\eta}_{Y_{ij}}^{L1})}{\mathbf{A}_{Z_i}^{L1}\mathbf{A}_{X_i}^{L1}\mathbf{R}_{X_i}^{L1}\mathbf{A}_{Z_i}^{L1}\mathbf{A}_{X_i}^{L1}\mathbf{A}_{Y_{ij}}^{L1}\mathbf{R}_{Y_{ij}}^{L1}\mathbf{A}_{Y_{ij}}^{L1}} \quad (19)$$

with latent variable interaction variance found with

$$\text{var}(\eta_{Z_i}^{L1}\eta_{X_i}^{L1}) = \text{var}(\eta_{Z_i}^{L1})\text{var}(\eta_{X_i}^{L1}) + \text{cov}(\eta_{Z_i}^{L1}, \eta_{X_i}^{L1})^2. \quad (20)$$

Here all components of the covariance are located at the within level eliminating the need for any consideration of unreliability of indicator cluster means ($\mathbf{R}^{L2}$). In fact, this correction operates as a correction for a single-level latent interaction (see Cox & Kelcey, 2021).

Results for the simulation study examining this correction can be found in Table 4 and indicate Croon's estimation performs well when estimating MLSEMs with an A1 moderation effect. Convergence failure rates were <4% at the smallest sample size condition and around 1% or less as n_2 exceeded 30. Bias and efficiency metrics for Croon's estimation again suggest it is an effective estimator and it continued to produce less biased estimates than the FS approach. The takeaway is that the newly derived corrections are accurate for latent interactions occurring at the within level.

3.5. Between Part of Level-one Moderator $\times$ within Part of Level-one Predictor (A2)

The final two corrections involve covariances with cross-level latent interactions. First, we present the correction for an A2 moderation effect when the covariance includes the between part of a level-one moderator (Z_j) and within part of a level-one predictor (X_i) and the outcome (Y_{ij}^{L1} ; see Figure 2e). This correction can be expressed as

$$\text{cov}(\eta_{Z_j}^{L2}\eta_{X_i}^{L1}, \eta_{Y_{ij}}^{L1}) = \frac{\text{cov}(\hat{\eta}_{Z_j}^{L2}\hat{\eta}_{X_i}^{L1}, \hat{\eta}_{Y_{ij}}^{L1})}{\mathbf{A}_{Z_j}^{L2}\mathbf{R}_{Z_j}^{L2}\mathbf{A}_{X_i}^{L1}\mathbf{A}_{Z_j}^{L2}\mathbf{A}_{X_i}^{L1}\mathbf{A}_{Y_{ij}}^{L1}\mathbf{R}_{Y_{ij}}^{L1}\mathbf{A}_{Y_{ij}}^{L1}} \quad (21)$$

with the latent interaction variance such that

$$\text{var}(\eta_{Z_j}^{L2}\eta_{X_i}^{L1}) = \text{var}(\eta_{Z_j}^{L2})\text{var}(\eta_{X_i}^{L1}) + 0^2. \quad (22)$$

The cross-level nature of the latent interaction does not change the general structure of the covariance correction (i.e., estimated covariance divided by components that capture unreliability). The estimated variance of the latent interaction does reflect a minor change stemming from the consideration of a cross-level interaction. The within- and between-level components are unrelated by design so the covariance term in the variance formula is essentially dropped. We include the 0^2 term as a placeholder to

Table 4. Summary of estimator performance with latent interaction A1 in terms of convergence failure rate, bias, and efficiency across sample size conditions.

Estimator	n_1	n_2	Fail converge	Bias		SD		RMSE	
				L2	L1	L2	L1	L2	L1
Croon	50	30	0.02	0.01	0.005	0.099	0.041	0.1	0.041
		60	0.014	0.006	0.005	0.066	0.03	0.067	0.03
		90	0.004	0.006	0.004	0.058	0.023	0.058	0.024
	20	30	0.036	0.009	0.01	0.103	0.065	0.104	0.066
		60	0.014	0.009	0.01	0.072	0.046	0.072	0.048
		90	0.01	0.002	0.008	0.059	0.037	0.059	0.038
FS	50	30	0.006	0.031	0.026	0.088	0.031	0.094	0.04
		60	0	0.031	0.026	0.057	0.022	0.065	0.034
		90	0.002	0.033	0.025	0.049	0.018	0.059	0.031
	20	30	0.006	0.037	0.028	0.089	0.051	0.096	0.058
		60	0.002	0.038	0.027	0.06	0.036	0.071	0.045
		90	0.002	0.036	0.025	0.047	0.028	0.059	0.038

Notes: Fail converge represents the rate estimators failed to converge. Estimators are abbreviated with Croon for Croon's bias-corrected estimation and FS for the uncorrected factor score regression approach.

Table 5. Summary of estimator performance with latent interaction A2 in terms of convergence failure rate, bias, and efficiency across sample size conditions.

Estimator	n_1	n_2	Fail converge	Bias		SD		RMSE	
				L2	L1	L2	L1	L2	L1
Croon	50	30	0.022	0.009	0.004	0.107	0.047	0.108	0.047
		60	0.002	0.007	0.004	0.071	0.033	0.072	0.033
		90	0.002	0.005	0.004	0.057	0.027	0.057	0.027
	20	30	0.034	0.01	0.009	0.114	0.065	0.114	0.066
		60	0.002	0.008	0.008	0.076	0.048	0.077	0.049
		90	0.016	0.006	0.007	0.061	0.038	0.062	0.039
FS	50	30	0	0.047	0.03	1.443	0.035	1.443	0.047
		60	0	0.03	0.03	0.062	0.024	0.069	0.04
		90	0	0.03	0.03	0.048	0.02	0.057	0.037
	20	30	0.008	0.034	0.032	0.099	0.05	0.104	0.06
		60	0	0.035	0.032	0.064	0.037	0.073	0.05
		90	0.002	0.036	0.031	0.051	0.029	0.062	0.043

Notes: Fail converge represents the rate estimators failed to converge. Estimators are abbreviated with Croon for Croon's bias-corrected estimation and FS for the uncorrected factor score regression approach.

illustrate this point. Results for the simulation study examining this correction and performance of Croon's estimation can be found in Table 5. Croon's performance in terms of convergence failure, bias, and efficiency paralleled previous simulation study results. The performance of the FS approach and the performance of Croon's relative to FS were also similar to previous results.

3.6. Level-two Moderator $\times$ within Part of Level-one Predictor (B1)

Our final correction involves the covariance of a B1 moderation effect that includes a level-two moderator (Z_j) and the within part of a level-one predictor (X_i ; see Figure 2f) expressed as

$$\text{cov}(\eta_{Z_j} \eta_{X_i}^{L1}, \eta_Y^{L1}) = \frac{\text{cov}(\hat{\eta}_{Z_j}^{L1}, \hat{\eta}_{Y_j}^{L1})}{\mathbf{A}_{Z_j}^{L1} \mathbf{A}_{X_i}^{L1} \mathbf{A}_{Z_j}^{L1} \mathbf{A}_{X_i}^{L1} \mathbf{A}_{Y_j}^{L1} \mathbf{A}_{Y_j}^{L1}} \quad (23)$$

with the variance of the latent interaction found using

$$\text{var}(\eta_{Z_j} \eta_{X_i}^{L1}) = \text{var}(\eta_{Z_j}) \text{var}(\eta_{X_i}^{L1}) + 0^2. \quad (24)$$

Table 6. Summary of estimator performance with latent interaction B1 in terms of convergence failure rate, bias, and efficiency across sample size conditions.

Estimator	n_1	n_2	Fail converge	Bias		SD		RMSE	
				L2	L1	L2	L1	L2	L1
Croon	50	30	0.026	0.011	0.012	0.105	0.056	0.106	0.058
		60	0.016	0.009	0.009	0.073	0.041	0.074	0.042
		90	0.006	0.005	0.007	0.06	0.034	0.06	0.035
	20	30	0.032	0.018	0.017	0.111	0.08	0.113	0.082
		60	0.014	0.01	0.014	0.075	0.055	0.076	0.057
		90	0.006	0.007	0.011	0.061	0.044	0.062	0.045
FS	50	30	0.002	0.031	0.048	0.094	0.045	0.099	0.066
		60	0.002	0.033	0.046	0.065	0.033	0.073	0.057
		90	0.002	0.031	0.045	0.05	0.027	0.059	0.053
	20	30	0.004	0.041	0.049	0.1	0.065	0.109	0.081
		60	0.002	0.038	0.047	0.062	0.045	0.073	0.065
		90	0.002	0.038	0.045	0.05	0.035	0.062	0.057

Notes: Fail converge represents the rate estimators failed to converge. Estimators are abbreviated with Croon for Croon's bias-corrected estimation and FS for the uncorrected factor score regression approach.

The correction and variance term for the B1 moderation effect is similar to the other cross-level interaction (see Equations (21) and (22)). The only noteworthy difference is that the components of a B1 moderation effect are measured at the corresponding level (e.g., within- or between level) so no $\mathbf{R}^{L2}$ term is necessary. Results for this final simulation study are presented in Table 6. Croon's performance in terms of convergence failure, bias, and efficiency paralleled previous simulation study results especially those of the other cross-level interaction (see results for A2 in Table 5). The takeaway from the simulation studies involving B1 and A2 moderation effects is that the newly derived corrections are not only accurate for models with latent interactions at the within- and between level but also for models with latent interactions occurring across levels.

3.7. Summary

Having presented the six corrections necessary for Croon's estimation to be utilized in with two-level SEMs that include latent interactions, we summarize the key results from the corresponding simulation studies. (1) The totality of results demonstrates the accuracy of the covariance (and variance) correction formulas involving two-level latent interactions. Specific results for Croon's estimation of the coefficients associated with each latent interaction are summarized in Table 7. These specific results parallel Croon's overall performance. Of note, relative bias for estimates of interaction term coefficients are near or below conventional guidelines of <0.05 (Hoogland & Boomsma, 1998) in almost all cases when $n_2 \geq 60$.

(2) Overall Croon's estimation convergence failure rates, bias, SDs of coefficient estimates, and RMSE values indicate it is a viable estimator for two-level SEMs with latent interactions; (3) Bias, SDs of coefficient estimates, and RMSE were fairly consistent across models with different types of latent interactions and their respective analytic models. These metrics generally decreased as n_2 , n_1 , or both increased and were generally smaller at the within level (L1) when compared to the between level (L2). Croon's estimation did have elevated levels of relative bias (e.g., ≈ 0.15)

Table 7. Croon estimation performance with specific latent interaction coefficients in terms of bias and efficiency across sample size conditions.

Type of interaction	n_2	Bias	Relative bias	SD
A3	30	0.024	0.119	0.112
	60	0.013	0.064	0.078
	90	0.009	0.045	0.065
C	30	0.030	0.154	0.142
	60	0.023	0.116	0.097
	90	0.018	0.089	0.076
B2 or D	30	0.029	0.142	0.139
	60	0.010	0.051	0.097
	90	0.009	0.045	0.073
A1	30	0.009	0.045	0.036
	60	0.008	0.040	0.025
	90	0.007	0.035	0.020
A2	30	0.041	0.03	0.036
	60	0.038	0.035	0.024
	90	0.038	0.03	0.019
B1	30	0.019	0.096	0.056
	60	0.011	0.055	0.044
	90	0.009	0.045	0.036

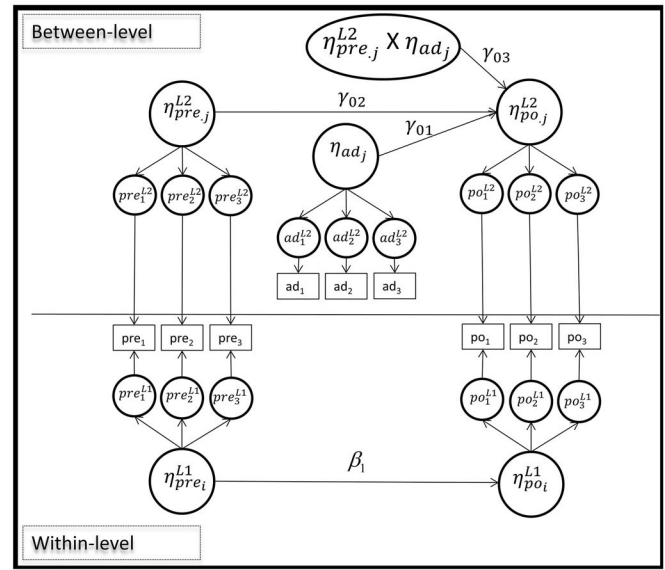
Note: $n_1 = 50$.

when cluster sample size (n_2) was less than 60; (4) Convergence failure rate was slightly higher for Croon's estimation when estimating models with latent interactions at the between level or across levels. That said, these convergence failure rates were almost always $\leq 5\%$ with rates $>1\%$ almost always occurring with $n_2 \leq 30$.

4. Illustration and Additional Considerations

In this section we use simulated data based on Capin et al. (2022) to illustrate Croon's estimation with multilevel latent interactions and consider the effect of scale reliability and intraclass correlation of items on estimator performance. Capin et al. (2022), examined the effect of the Promoting Adolescents' Comprehension of Text (PACT) reading program on academic outcomes in eighth grade students. The study utilized a MLSEM with students (within level) or level-one) nested within classes (between level or level-two) to consider relations among several latent variables including instructional quality, treatment adherence, student social studies content knowledge (pre- and post-test scores), and student reading comprehension (pre- and post-test scores).

Our adapted research question inspired by Capin et al. (2022) includes a focal predictor of adherence to the PACT program (adherence) measured at level-two or the class-level, a student's initial reading comprehension level as a moderator (reading pre-test) measured at level-one or the student-level, and the student's reading comprehension level after receiving PACT program instruction (reading post-test) as the outcome measured at level-one or the student-level (see Figure 3). This examination would have a particular interest in the main effect of the PACT program on student's reading post-test but also includes a multilevel latent interaction. Specifically, it includes a type D multilevel latent interaction ($1 \times (2 \rightarrow 1)$) that involves the between component of the reading pre-test moderating the relationship between the level-two PACT program adherence predictor and reading post-test outcome. In other words, the initial reading comprehension ability of a teacher's students may

**Figure 3.** Conceptual model for illustration MLSEM with latent interaction composed of the between part of a level-one moderator and a level-two predictor (D).

moderate the effect of the PACT program adherence. Perhaps the program is more beneficial with groups of students that on average start with lower levels of reading comprehension or it may be less beneficial under these conditions. The inclusion of the interaction term allows us to consider these possibilities.

Following the steps for Croon's estimation, a factor model would be fit using maximum likelihood estimation and factor scores generated for all latent variables (e.g., adherence, pre-test, post-test), the variance-covariance matrix of these factor scores would be estimated, then the Croon's correction would be applied before estimated the structural model with this corrected variance-covariance matrix (see R code for data generation and analysis code). Kelcey et al. (2021) provides details for Croon's estimation that corrects all variances and covariances involving the latent variables while this work provides the correction for covariances involving the latent interaction term (e.g., pre-test X adherence). In technical terms, the correction for the covariance involving the between part of a level-one latent moderator and a level-two latent predictor and $(\text{cov}(\eta_{Z_j}^{L2}, \eta_{X_j}^{L2}))$ is

$$\text{cov}(\eta_{Z_j}^{L2}, \eta_{Y_j}^{L2}) = \frac{\text{cov}(\hat{\eta}_{Z_j}^{L2}, \hat{\eta}_{Y_j}^{L2})}{\mathbf{A}_{Z_j}^{L2} \mathbf{R}_{Z_j}^{L2} \mathbf{A}_{X_j}^{L2} \mathbf{A}_{X_j}^{L2} \mathbf{A}_{Y_j}^{L2} \mathbf{R}_{Y_j}^{L2} \mathbf{A}_{Y_j}^{L2}} \quad (25)$$

with variance of the interaction such that

$$\text{var}(\eta_{Z_j}^{L2} \eta_{X_j}^{L2}) = \text{var}(\eta_{Z_j}^{L2}) \text{var}(\eta_{X_j}^{L2}) + \text{cov}(\eta_{Z_j}^{L2}, \eta_{X_j}^{L2})^2. \quad (26)$$

With a corrected variance-covariance matrix, the structural model can be estimated using maximum likelihood estimation.

Using simulated data, we estimate the illustrative example model. For simplicity, we set the conditions in the example to match those described for the B2 multilevel moderation effect simulation study (e.g., measurement models) but used

Table 8. Estimator performance for illustration example with latent interaction D in terms of convergence failure rate, bias, and efficiency.

Estimator	Fail converge	RMSEA		β_1		γ_{01}		γ_{02}		γ_{03}	
		L2	L1	Est.	SD	Est.	SD	Est.	SD	Est.	SD
Croon	0.02	0.1	0.056	0.181	0.053	0.189	0.101	0.2	0.098	0.185	0.098
FS	0.002	0.088	0.065	0.149	0.043	0.162	0.05	0.176	0.077	0.165	0.078

Note: All coefficients were simulated with values of 0.2 and $n_2 = 47$ with $n_1 = 16$.

Table 9. Summary of estimator performance with latent interaction D in terms of convergence failure rate, bias, and efficiency across sample size conditions and different variance decompositions for the multilevel factor models ($\epsilon_{Z_j}^{L2}$, $\epsilon_{Z_j}^{L2}$, and $\epsilon_{Y_j}^{L2}$).

Item ICC ($\epsilon_{Z_j}^{L2}$, $\epsilon_{Y_j}^{L2}$)	Estimator	n_1	n_2	Fail converge	Bias		SD		RMSE	
					L2	L1	L2	L1	L2	L1
0.05	Croon	50	30	0.066	0.017	0.003	0.113	0.042	0.114	0.042
			60	0.012	0.007	0.004	0.077	0.029	0.078	0.029
			90	0.016	0.007	0.002	0.063	0.024	0.064	0.024
		20	30	0.226	0.011	0.01	0.126	0.061	0.126	0.062
			60	0.154	0.002	0.009	0.082	0.044	0.082	0.045
			90	0.116	0.005	0.005	0.066	0.037	0.066	0.037
	FS	50	30	0.006	0.016	0.049	0.103	0.032	0.105	0.058
			60	0.002	0.013	0.051	0.066	0.022	0.068	0.055
			90	0.002	0.016	0.049	0.054	0.018	0.057	0.053
		20	30	0.09	0.019	0.052	0.105	0.047	0.107	0.07
			60	0.056	0.017	0.052	0.066	0.034	0.068	0.062
			90	0.034	0.017	0.049	0.055	0.028	0.058	0.057
0.4	Croon	50	30	0.032	0.023	0.005	0.143	0.036	0.145	0.036
			60	0.004	0.02	0.005	0.096	0.026	0.098	0.027
			90	0	0.011	0.004	0.078	0.022	0.079	0.023
		20	30	0.02	0.025	0.008	0.146	0.055	0.149	0.056
			60	0.008	0.017	0.011	0.101	0.042	0.103	0.043
			90	0.004	0.015	0.008	0.079	0.034	0.08	0.035
	FS	50	30	0.006	0.044	0.035	0.116	0.031	0.124	0.046
			60	0	0.048	0.035	0.077	0.022	0.09	0.041
			90	0	0.044	0.035	0.06	0.019	0.075	0.039
		20	30	0.002	0.048	0.034	0.121	0.048	0.131	0.058
			60	0.002	0.049	0.036	0.076	0.036	0.091	0.051
			90	0	0.051	0.034	0.058	0.029	0.078	0.045

Note: Fail converge represents the rate estimators failed to converge. Estimators are abbreviated with Croon for Croon's bias-corrected estimation and FS for the uncorrected factor score regression approach.

a sample size that reflected Capin et al. (2022) such that the cluster or class sample size was 47 ($n_2 = 47$) and individual per cluster or student per class sample size was 16 ($n_1 = 16$). Let us first consider specific coefficient estimates across simulated data sets (see Table 8). Across the 500 replications, Croon's estimation mostly avoided convergence failure issues even with the limited sample sizes. Croon's estimation did slightly underestimate the relationships between program adherence, the within-level component of pre-test scores, and the interaction term with the post-test outcome but was more accurate than the factor score approach. Of particular importance is γ_{03} which captures the magnitude of the interaction. Croon's estimate of 0.185 indicates relative bias of 0.075, just exceeding typical thresholds but substantially outperforming the factor score approach. When considering both accuracy and efficiency (i.e., RMSEA), Croon's estimation outperformed or was performed similarly to the factor score approach. Overall, Croon's approach demonstrated apt performance with the hypothetical example from Capin et al. (2022) and this illustration serves as a showcase for possible applications of the estimator in substantive research.

We also use this example as the context to conduct an initial evaluation of the influence of scale reliability and

item intraclass correlation on the accuracy and sensitivity of the Croon's method in multilevel settings. We again set simulation conditions in the example to match those described for the B2 multilevel moderation effect simulation study but consider two different intraclass correlation coefficients of the indicator error terms ($\epsilon_{Z_j}^{L2}$, $\epsilon_{Y_j}^{L2}$) in the multilevel factor models (see Table 9) and two scale reliability conditions by varying item factor loadings (see Table 10). While these conditions only provide a preliminary and limited assessment, we can glean some important implications.

First, as indicator error term ICCs approached zero convergence failure rate for Croon's estimation, and to a lesser extent the factor score approach, increased. The influence of indicator error term ICC on estimator bias and efficiency was minimal. In terms of our example, if very little indicator error variance was attributable to the class-level then estimation would be slightly more difficult. Second, increases to factor loadings (i.e., increased scale reliability) decreased the convergence failure rate of Croon's and also produced slight decreases to bias while improving efficiency. In our example, improvements in measuring PACT program adherence and reading comprehension (pre- and post-test) would improve model estimation. It should be noted that these results echo previous

Table 10. Summary of estimator performance with latent interaction D in terms of convergence failure rate, bias, and efficiency across sample size conditions and different scale reliability.

Factor loadings	Estimator	n_1	n_2	Fail converge	Bias		SD		RMSE	
					L2	L1	L2	L1	L2	L1
0.5	Croon	50	30	0.104	0.02	0.005	0.195	0.056	0.196	0.057
			60	0.016	0.012	0.006	0.122	0.039	0.123	0.04
			90	0.004	0.011	0.003	0.091	0.032	0.092	0.032
		20	30	0.07	0.039	0.015	0.179	0.088	0.184	0.089
			60	0.012	0.022	0.009	0.123	0.06	0.125	0.061
			90	0.008	0.013	0.008	0.099	0.048	0.101	0.048
	FS	50	30	0.008	0.05	0.078	0.145	0.037	0.154	0.087
			60	0	0.049	0.079	0.094	0.026	0.106	0.083
			90	0	0.054	0.078	0.07	0.021	0.088	0.08
		20	30	0.002	0.101	0.078	0.752	0.063	0.762	0.1
			60	0	0.065	0.077	0.09	0.041	0.111	0.087
			90	0.002	0.061	0.076	0.074	0.032	0.096	0.083
1.5	Croon	50	30	0.004	0.01	0.004	0.121	0.034	0.122	0.034
			60	0	0.005	0.004	0.077	0.025	0.077	0.025
			90	0	0.003	0.003	0.064	0.019	0.064	0.019
		20	30	0.022	0.01	0.01	0.133	0.055	0.133	0.056
			60	0.002	0.004	0.008	0.085	0.04	0.085	0.04
			90	0	0.003	0.01	0.068	0.032	0.068	0.033
	FS	50	30	0	0.02	0.03	0.105	0.029	0.107	0.042
			60	0	0.024	0.03	0.065	0.021	0.069	0.036
			90	0	0.02	0.029	0.055	0.016	0.059	0.033
		20	30	0.002	0.031	0.03	0.104	0.05	0.109	0.058
			60	0	0.029	0.028	0.069	0.036	0.075	0.045
			90	0	0.032	0.03	0.053	0.029	0.062	0.041

Notes: The first factor loading of each indicator was fixed at 1.0 to set the scale of the latent variable all other factor loadings were set at 1.5 and 0.5 in each condition, respectively. Fail Converge represents the rate estimators failed to converge. Estimators are abbreviated with Croon for Croon's bias-corrected estimation and FS for the uncorrected factor score regression approach.

findings related to Croon's estimation (e.g., Kelcey et al., 2021).

5. Discussion

Estimation of MLSEMs with latent interactions, especially with limited samples sizes, is often not feasible with typical estimators (e.g., maximum likelihood) because of its computational complexity. Croon's estimation has performed well in a variety of SEMs but was not yet available for MLSEMs with latent interactions. The purpose of this study was (a) to extend Croon's estimation to accommodate two-level SEMs with latent interactions and (b) assess the accuracy of new Croon's based corrections and the performance of Croon's estimation with a variety of multilevel moderation models. Following the multilevel moderation effect conceptualizations presented in Preacher et al. (2016), we identified and developed six covariance and variance corrections necessary for Croon's estimation to accommodate all two-level moderation effects. We then conducted simulation studies examining each of the corrections to demonstrate their accuracy, the feasibility of Croon's estimation to estimate MLSEMs with latent interactions, and evaluate the comparative performance of Croon's against uncorrected factor score regression. Results indicated that the covariance corrections were accurate and Croon's estimation is a viable estimator for two-level SEMs with latent interactions.

The primary contributions of this work are (1) the Croon's corrections necessary for estimating MLSEMs with latent interactions and (2) an initial understanding of how

well Croon's estimation performs with MLSEMs that include latent interactions. The availability of Croon's estimation for MLSEMs with latent interactions expands the capacity of researchers to consider complex theories in multilevel settings. In these settings, we found typical maximum likelihood estimation to be impractical as model complexity increased and FS approaches biased. The struggles of these estimators with limited sample sizes makes the availability of alternative estimators such as Croon's crucial. Simulation study results increase understanding of the sample sizes requirements for Croon's estimation to dependably produce accurate and efficient estimates for different multilevel moderation models. Our results suggested that Croon's estimation can deliver dependable estimates with sample sizes as small as $n_2 = 30$ and $n_1 = 20$ whereas maximum likelihood estimation was not feasible with similar sample sizes and uncorrected factor score regression estimates were biased.

There are limitations to this work. We considered multilevel moderation models with all latent variables measured using three indicators per latent variable. We did not vary measurement models or their conditions. While the complexity of the latent variable measurement models is novel, Croon's estimation is not limited to these factor models. We recommend future research consider different measurement model conditions including various levels of composite reliability, correlated measurement residuals, and factor loadings that vary across levels (i.e., cross-level non-invariance). Structural model modifications that were not considered here but may be of interest include more complex path models, latent moderated mediation, and corrections that

accommodate random slopes. Another pertinent area of research is the robustness of Croon's estimation when assumptions related to the measurement and structural model are not met (e.g., Blinded, 2022; Robitzsch, 2022). Finally, we recommend a more comprehensive examination of the factors influencing estimation of MLSEMs with latent interactions including a wider variety of n_2 and n_1 conditions, different decompositions of outcome variance across levels, and a comparison Croon's estimation to Bayesian approaches. These limitations suggest many details surrounding Croon's estimation of MLSEMs with latent interactions need to be investigated further but we are confident in recommending its use even with limited sample sizes.

Funding

This article is based on work funded by the National Science Foundation [#1552535 and # 1760884]. The opinions expressed herein are those of the authors and not the funding agency.

ORCID

Kyle Cox  <http://orcid.org/0000-0002-7173-4701>

References

- Aiken, L., West, S., & Reno, R. (1991). *Multiple regression: Testing and interpreting interactions*. Sage Publications.
- Asparouhov, T., & Muthén, B. (2021). Bayesian estimation of single and multilevel models with latent variable interactions. *Structural Equation Modeling: A Multidisciplinary Journal*, 28, 314–328. <https://doi.org/10.1080/10705511.2020.1761808>
- Authors Redacted. (in press). Croon's bias corrected estimation for multilevel structural equation models with non-normal indicators and model misspecifications. *Educational and Psychological Measurement*.
- Bohrnstedt, G., & Goldberger, A. (1969). On the exact covariance of products of random variables. *Journal of the American Statistical Association*, 64, 1439–1442. <https://doi.org/10.1080/01621459.1969.10501069>
- Brandt, H., Umbach, N., Kelava, A., & Bollen, K. (2020). Comparing estimators for latent interaction models under structural and distributional misspecifications. *Psychological Methods*, 25, 321–345. <https://doi.org/10.1037/met0000231>
- Capin, P., Roberts, G., Clemens, N. H., & Vaughn, S. (2022). When treatment adherence matters: Interactions among treatment adherence, instructional quality, and student characteristics on reading outcomes. *Reading Research Quarterly*, 57, 753–774. <https://doi.org/10.1002/rrq.442>
- Cox, K., & Kelcey, B. (2021). Croon's bias corrected estimation of latent interactions. *Structural Equation Modeling: A Multidisciplinary Journal*, 28, 863–874. <https://doi.org/10.1080/10705511.2021.1922283>
- Cox, K., & Kelcey, B. (2022). Croon's bias corrected estimation for multilevel structural equation models with non-normal indicators and model misspecifications. *Educational and Psychological Measurement*, 001316442210804.
- Croon, M. (2002). Using predicted latent scores in general latent structure models. In G. Marcoulides & I. Moustaki (Eds.), *Latent variable and latent structure modeling* (pp. 195–223). Erlbaum.
- Devlieger, I., Mayer, A., & Rosseel, Y. (2016). Hypothesis testing using factor score regression: A comparison of four methods. *Educational and Psychological Measurement*, 76, 741–770. <https://doi.org/10.1177/0013164415607618>
- Devlieger, I., & Rosseel, Y. (2017). Factor score path analysis: An alternative for SEM? *Methodology*, 13, 31–38. <https://doi.org/10.1027/1614-2241/a000130>
- Devlieger, I., & Rosseel, Y. (2020). Multilevel factor score regression. *Multivariate Behavioral Research*, 55, 600–624.
- Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 25, 621–638. <https://doi.org/10.1080/10705511.2017.1402334>
- Hayes, T., & Usami, S. (2020a). Factor score regression in connected measurement models containing cross-loadings. *Structural Equation Modeling: A Multidisciplinary Journal*, 27, 942–951. <https://doi.org/10.1080/10705511.2020.1729160>
- Hayes, T., & Usami, S. (2020b). Factor score regression in the presence of correlated unique factors. *Educational and Psychological Measurement*, 80, 5–40. <https://doi.org/10.1177/0013164419854492>
- Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, 26, 329–367. <https://doi.org/10.1177/0049124198026003003>
- Hoshino, T., & Bentler, P. (2013). Bias in factor score regression and a simple solution. In A. R. de Leon & K. C. Chough (Eds.), *Analysis of mixed data: Methods & applications* (pp. 43–62). Chapman and Hall/CRC.
- Kelcey, B. (2019). A robust alternative estimator for small to moderate sample SEM: Bias-corrected factor score path analysis. *Addictive Behaviors*, 94, 83–98. <https://doi.org/10.1016/j.addbeh.2018.10.032>
- Kelcey, B., Cox, K., & Dong, N. (2021). Croon's bias-corrected factor score path analysis for small- to moderate-sample multilevel structural equation models. *Organizational Research Methods*, 24, 55–77. <https://doi.org/10.1177/1094428119879758>
- Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, 96, 201–210. <https://doi.org/10.1037/0033-2909.96.1.201>
- Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, 65, 457–474. <https://doi.org/10.1007/BF02296338>
- Loncke, J., Eichelshheim, V. I., Branje, S. J. T., Buyse, A., Meeus, W. H. J., & Loeys, T. (2018). Factor score regression with social relations model components: A case study exploring antecedents and consequences of perceived support in families. *Frontiers in Psychology*, 9, 1699–1619. <https://doi.org/10.3389/fpsyg.2018.01699>
- Muthén, L. K., & Muthén, B. O. (1998–2021). *Mplus user's guide* (8th ed.). Muthén & Muthén.
- Ng, J., & Chan, W. (2020). Latent moderation analysis: A factor score approach. *Structural Equation Modeling: A Multidisciplinary Journal*, 27, 629–648. <https://doi.org/10.1080/10705511.2019.1664304>
- Preacher, K., Zhang, Z., Zyphur, M., & Preacher, K. (2016). Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychological Methods*, 21, 189–205.
- R Development Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Robitzsch, A. (2022). Comparing the robustness of the structural after measurement (SAM) approach to structural equation modeling (SEM) against local model misspecifications with alternative estimation approaches. *Stats*, 5, 631–672. <https://doi.org/10.3390/stats5030039>
- Rosseel, Y. (2020). Small sample solutions for structural equation modeling. In R. Van de Schoot and M. Miočević (Eds.), *Small sample size solutions: A how to guide for applied researchers and practitioners*. Routledge.
- Rosseel, Y., & Loh, W. W. (in press). A structural after measurement (SAM) approach to structural equation modeling. *Psychological Methods*. <https://osf.io/pekbm/>
- Schoemann, A., & Jorgensen, T. (2021). Testing and interpreting latent variable interactions using the semTools package. *Psych*, 3, 322–335. <https://doi.org/10.3390/psych3030024>
- Skrondal, A., & Laake, P. (2001). Regression among factor scores. *Psychometrika*, 66, 563–575. <https://doi.org/10.1007/BF02296196>

Smid, M., McNeish, D., Miočević, M., & van de Schoot, R. (2020). Bayesian versus frequentist estimation for structural equation models in small sample contexts: A systematic review. *Structural Equation Modeling: A Multidisciplinary Journal*, 27, 131–161. <https://doi.org/10.1080/10705511.2019.1577140>

Smid, S., & Rosseel, Y. (2020). SEM with small samples: Two-step modeling and factor score regression versus Bayesian estimation with informative priors. In *Small sample size solutions* (1st ed., pp. 239–254). Routledge.

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

Zitzmann, S., & Helm, C. (2021). Multilevel analysis of mediation, moderation, and nonlinear effects in small samples, using expected a posteriori estimates of factor scores. *Structural Equation Modeling: A Multidisciplinary Journal*, 28, 529–546. <https://doi.org/10.1080/10705511.2020.1855076>

PROOF ONLY