Estimating latent baseline-by-treatment interactions in statistical mediation analysis

Oscar Gonzalez[1], **Jeno R. Millechek**[1], **A. R. Georgeson**[2]

[1]University of North Carolina at Chapel Hill

[2]Arizona State University

Abstract

Statistical mediation analysis is used to uncover intermediate variables, known as mediators [M], that explain how a treatment [X] changes an outcome [Y]. Often, researchers examine whether baseline levels of M and Y moderate the effect of X on posttest M or Y. However, there is limited guidance on how to estimate baseline-by-treatment interaction (BTI) effects when M and Y are latent variables, which entails the estimation of latent interaction effects. In this paper, we discuss two general approaches for estimating latent BTI effects in mediation analysis: using structural models or scoring latent variables prior to estimating observed BTIs and correcting for unreliability. We present simulation results describing bias, power, type 1 error rates, and interval coverage of the latent BTIs and mediated effects estimated using these approaches. These methods are also illustrated with an applied example. R and M*plus* syntax are provided to facilitate the implementation of these approaches.

In the social sciences, statistical mediation analysis is used to investigate the intermediate variables, known as mediators (M), by which a treatment (X) led to a change in an outcome (Y; MacKinnon, 2008). Recent examples of mediation hypotheses include a program on adaptive regulation strategies (X) decreasing externalizing problems (Y) by improving emotion regulation (M) in adolescents (te Brinke et al., 2021), and an imagery rescripting treatment (X) decreasing nightmare distress (Y) by increasing self-efficacy (M) in individuals with nightmare disorder (Kunze et al., 2019). Often, researchers include baseline levels of M and Y in the mediation model (i.e., measures prior to the treatment or intervention) to account for confounding of the M to Y relation at posttest (Valente et al., 2020). Furthermore, with the inclusion of baseline measures, researchers could also test whether the effect of the program depends on the baseline levels of M and Y (Baron & Kenny, 1986; Morgan-Lopez & MacKinnon, 2006), which helps inform future implementation of prevention programs (MacKinnon, 2008). For example, in a randomized prevention program (X) to decrease the use of anabolic steroids (Y) in high school athletes by increasing their knowledge of the effects of steroids on health (M), the program was shown to be most effective for athletes who had high intentions of using steroids at baseline (Fritz et al., 2005). As such, baseline levels of M and Y could moderate the effect of the program on

corresponding author: ogonza13@unc.edu.

the mediator or the outcome, and this moderation can be assessed by including baseline-by-treatment interactions in the model (Morgan-Lopez & MacKinnon, 2006).

Measures of M and Y, such as sum or mean scores, often contain measurement error, which, if ignored, might either inflate or attenuate parameter estimates (Gonzalez & MacKinnon, 2021; Hoyle & Kenny, 1999; le Cessie et al., 2012; VanderWeele et al., 2012). Measurement error can also reduce the reliability of the interaction terms (i.e., XM, XY), which attenuates both the interaction effects in regression (Aiken & West, 1991; Bohrnstedt & Marwell, 1978; Busemeyer & Jones, 1983; Kenny & Judd, 1984; McClelland & Judd, 1993) and mediation models (Cheung & Lau, 2017; Gonzalez & Valente, in press). A common solution for addressing measurement error is to specify multiple-indicator latent variables for M and Y. If M or Y interact with X, latent interactions must be estimated. There are many methods for estimating this type of interaction effects (e.g., Asparouhov & Muthen, 2021; Cox & Kelcey, 2021; Hsiao et al., 2018; Klein & Moosbrugger, 2000; Marsh et al., 2012; Wall & Amemiya, 2003).

Previous studies suggest that mediation effects can be accurately estimated in single mediator models with one latent interaction using methods cited above (e.g., Cheung & Lau, 2017; Gonzalez & Valente, in press). However, it is unclear if those same findings hold in two-wave mediation models with latent baseline-by-treatment interaction (BTI) effects. Estimating interaction effects in mediation models with two waves of data is more complex than in mediation models with a single wave of data because the inclusion of latent M and Y at pretest and posttest creates two latent interaction terms (instead of one), and four interaction effects (instead of one). Additionally, correlated item residuals often arise when the same latent variable (e.g., M or Y) is assessed with the same measure at pretest and posttest, which must be adjusted for to obtain accurate parameter estimates. In this paper, we describe and evaluate the performance of methods developed to estimate latent interaction effects in two-wave mediation models with BTI effects and provide the following contributions. First, we provide covariance expectations for a better understanding of the two-wave mediation model with BTIs and apply the methods to empirical data. Second, we use Monte Carlo simulations to assess the recovery of the mediation paths and interaction effects across methods. Third, we provide code to estimate this model with the discussed methods. Our goal is to provide researchers with resources to understand and estimate in two-wave mediation models with latent BTIs and facilitate the implementation of these models.

The structure of the paper is the following. First, we introduce the statistical mediation model with BTIs. Then, we discuss five different ways to estimate latent BTI effects in the mediation model. Next, we present the results of a Monte Carlo simulation study on the bias, power, type 1 error rate, and interval coverage of the parameter estimates of the statistical mediation model with latent BTIs. Lastly, we demonstrate the methods with data from the JOBS II study. In the supplement, we include the covariance expectations for mediation models with BTIs, along with R and M*plus* syntax for estimating those models.

Two-wave mediation model

In this paper, we focus on the simplest mediation model for experimental studies with pretest and posttest data (Valente & MacKinnon, 2017). Although there are other ways to capture change in mediation models (e.g., Howe, 2019; Montoya, 2019; Valente & MacKinnon, 2017), the two-wave mediation model allows researchers to control for baseline levels of M and Y, which in turn mitigates confounding of the posttest M to Y relation (Valente et al., 2020). Two-wave mediation models (some with other covariates) have been used to evaluate interventions for drug use (e.g., MacKinnon, 2001), internalizing symptoms (e.g., Jensen et al., 2014; Perrino et al., 2014), depressive symptoms (e.g., Chaudoir et al., 2021), antisocial behavior (e.g., Smith et al., 2014), and substance use and school discipline (e.g., Gonzales et al., 2012).

A mediation model with a randomized treatment, continuous pretest and posttest measures, and BTIs can be expressed as:

$$M_2 = i_M + aX + s_m M_1 + c_y Y_1 + h_1 X M_1 + h_2 X Y_1 + e_M$$

(1)

$$Y_2 = i_Y + c'X + b M_2 + s_y Y_1 + c_m M_1 + h_3 X M_1 + h_4 X Y_1 + e_Y \ .$$

(2)

In Eq. 1 and 2, X represents a binary randomized treatment/control group indicator; M_1 and M_2 represent the mediator at pretest and posttest, respectively; and Y_1 and Y_2 represent the outcome at pretest and posttest, respectively. Note that in randomized treatments, X does not correlate with pretest measures M_1 and Y_1 in expectation. The a-path is the relation between X and M_2, s_m is the temporal stability of M, c_y is the cross-lag path between M_2 and Y_1, b is the relation between M_2 and Y_2, c' is the relation between X and Y_2, s_y is the temporal stability of Y, and c_m is the cross-lag path between Y_2 and M_1. Furthermore, the four coefficients h_1, h_2, h_3 and h_4 represent the BTIs, where the relation between X and M_2 or Y_2 is moderated by the baseline scores. In the absence of BTIs, the mediated effect is defined by the product of ab. We can construct confidence intervals and test for statistical significance of ab using the distribution of the product of two random variables (Tofighi & MacKinnon, 2011), bootstrapping (MacKinnon et al., 2004), or deriving a Bayesian credible interval for ab (Yuan & MacKinnon, 2009). If h_1 or h_2 are significant, then there is a baseline-by-treatment interaction on the mediator. As such, the a-path varies as a function of the pretest measures, indicating moderated mediation (Hayes, 2015; MacKinnon, 2008). If h_3 or h_4 are significant, then there is a baseline-by-treatment interaction on the outcome, so the c'-path varies as a function of the pretest measures.

To ascribe a causal interpretation to the mediated effect, some assumptions must hold. For example, we assume that we have specified the appropriate functional form and temporal precedence between the variables. We also assume that the scores on all variables have

high reliability and can be ascribed a valid interpretation (Gonzalez & MacKinnon, 2021). Furthermore, there are four no-confounding assumptions needed to identify the mediated effect (Imai et al., 2010; Pearl, 2001; Valeri & VanderWeele, 2013):

1. No confounders in the relation between X and M_2, after controlling for covariates.

2. No confounders in the relation between M_2 and Y_2, after controlling for covariates.

3. No confounders in the relation between X and Y_2, after controlling for covariates.

4. No confounders in the relation between M_2 and Y_2 that are affected by X.

In randomized treatment scenarios, assumptions 1 and 3 are typically satisfied, but assumptions 2 and 4 are not satisfied because individuals self-select their value on M_2. Controlling for pretest levels of M and Y makes meeting assumptions 2 and 4 more tenable because M_1 and Y_1 form a joint confounder of the M_2 to Y_2 relation (Mayer et al., 2014; Valente et al., 2020) via the model stabilities and the cross-lags, which we control for in the model. Sensitivity analyses can be used to determine how likely it is that assumptions 2 and 4 are met (Cox et al., 2013; Fritz et al., 2016; Liu & Wang, 2021). Lastly, as mentioned above, the model presented in Eq. 1 and Eq. 2 has linear relations, continuous M and Y, and no XM_2 interaction, so traditional mediated effects estimated from regression models and the causal mediation estimates from the potential outcomes framework are equivalent (MacKinnon et al., 2020; Valeri & VanderWeele, 2013).

Statistical mediation model with latent variables

As mentioned above, there are many situations in which the mediator or the outcome are measured with error (Gonzalez & MacKinnon, 2021). One way to adjust for measurement error is to specify latent variables for M and Y at each time point. The relation between items and latent variables can be represented using the linear factor model:

$$m_{1ij} = \tau_{m_1j} + \lambda_{m_1j}M_{1i} + e_{m_1ij} ; \quad m_{2ij} = \tau_{m_2j} + \lambda_{m_2j}M_{2i} + e_{m_2ij}$$

$$(3)$$

$$y_{1ij} = \tau_{y_1j} + \lambda_{y_1j}Y_{1i} + e_{y_1ij} ; \quad y_{2ij} = \tau_{y_2j} + \lambda_{y_2j}Y_{2i} + e_{y_2ij},$$

$$(4)$$

where M_{Ti} and Y_{Ti} represent the latent scores for individual i at time $= T$, which are assessed by a set of j items in vectors m_{Tij} and y_{Tij}. In this case, τ_{mTj} and τ_{yTj} represent the item intercepts, λ_{mTj} and λ_{yTj} are the factor loadings, and e_{mTij} and e_{yTij} are the unique scores for individual i on item j at time T. For our purposes, we assume longitudinal scalar measurement invariance between pretest and posttest measures in the mediation model (e.g., Georgeson et al., 2021),

$$\tau_{1j} = \tau_{2j} = \tau_j; \quad \lambda_{1j} = \lambda_{2j} = \lambda_j,$$

(5)

and estimate correlated residuals between like-indicators of the same construct across time. An example of the model is presented in Figure 1. Because the model now includes latent variables, the BTIs are necessarily latent interactions. Given that M_1, M_2, Y_1, and Y_2 are not observed, we consider two general approaches to estimate the latent BTI effect. The first approach is to compute an observed scores for M_1, M_2, Y_1, and Y_2, compute observed BTIs, and then adjust for unreliability and correlated errors. The second approach is to use structural models to estimate the latent BTI effects. We discuss five methods across these two general approaches next.

Scoring latent variables with reliability adjustment

As mentioned above, when measurement error in M and Y is ignored, interaction effects and mediation paths are attenuated (Gonzalez & Valente, in press). One can correct for measurement error by fitting the mediation model with BTIs to the full covariance matrix among the observed variables and interactions (e.g., the covariance matrix of X, M_1, M_2, Y_1, Y_2, XM_1, and XY_1), and manually correcting the variances and covariances of the scores on M and Y for unreliability and the effect of the correlated residuals. Unreliability inflates observed score variances (i.e., variances of observed M_1, M_2, Y_1, Y_2, XM_1, and XY_1) but not the covariances. Ignoring correlated residuals among indicators of two latent variables inflates the covariances of the observed scores that represent said latent variables (e.g., the covariance of observed M_1 and M_2; Cole & Preacher, 2014; Gonzalez & MacKinnon, 2021; Hayes & Usami, 2020). Here, we apply scoring methods to estimate latent interaction effects (Croon, 2002; Cox & Kelcey, 2021; Hayes & Usami, 2020) to the two-wave mediation model and extend the methods when noted.

The first step is to obtain scores for M and Y. One option is to score the latent variable is to use unit-weighted summed scores (**Method 1**, referred to as *summed scores with reliability adjustment*; MacKinnon, 2008). For example, summed score M_{1s} is calculated by adding the item responses of M_1. Another option is to use factor scores (**Method 2,** referred to as *factor scores with reliability adjustment*; Bartlett, 1937; Lawley & Maxwell, 1971; Skrondal & Laake, 2001). Factor scores can be thought of as weighted summed scores where the weights are determined by the parameter estimates from the linear factor model. In this paper, we use Bartlett scores for Method 2. The weights (i.e., scoring matrix) are estimated by using the expression,

$$B_w = \left(\Lambda' \Theta_e^{-1} \Lambda \right)^{-1} \Lambda' \Theta_e^{-1},$$

(6)

where Λ is a j x p matrix of factor loadings, where j are the indicators and p are the latent variables, and Θ_e is the j x j covariance matrix of the unique scores. To obtain the

B_w weights to score M_1 and M_2, Λ and Θ_e are obtained from a correlated two-factor model with correlated residuals, specifying longitudinal scalar invariance. Once obtained, the item responses are multiplied by the weights and added to estimate factor scores M_{1f} and M_{2f}. A similar procedure is used to estimate Y_{1f} and Y_{2f}.

The second step is to estimate the full covariance matrix of observed scores and interactions and correct specific submatrices of the full covariance matrix. For Method 2, we propose to apply the versions of Croon's (2002) correction to models with interaction effects (Cox & Kelcey, 2021) and correlated residuals (Hayes & Usami, 2020) simultaneously. For example, to correct the 2×2 submatrix Σ_{M_f}, comprised of the variances of M_{1f} and M_{2f} and their covariance, we need the scoring weights for M, B_{wM}, and the submatrix of the variances and covariances of the residuals for the j indicators of M, Θ_{eM}. As such, one can use: [1]

$$\Sigma_{M_{fc}} = \Sigma_{M_f} - B_{wM}\Theta_{eM}B_{wM}'$$

(7)

Similar steps are used to correct Σ_{Y_f}. To correct the submatrices in the full covariance matrix for Method 1, we adapt Eq. 8 to summed scores by replacing B_{wM} with an p x j summing matrix S_M. S_M is populated by 1s on the entries indicating which items load on each variable and 0s elsewhere. For example, to correct submatrix Σ_{M_s}, comprised of the variances of M_{1s} and M_{2s} and their covariance, one can use:[2]

$$\Sigma_{M_{sc}} = \Sigma_{M_s} - S_M\Theta_{eM}S_M'.$$

(8)

Similar steps are used to correct Σ_{Y_s}. Once $\Sigma_{M_{fc}}$, $\Sigma_{Y_{fc}}$, $\Sigma_{M_{sc}}$, and $\Sigma_{Y_{sc}}$ are obtained, then the variances of the observed interaction terms can be corrected. For interaction terms, their variance can be implied by the variances and covariances of their components (Bohrnstedt & Marwell, 1978). For example, when X and M_{1s} are centered, the implied variance of XM_{1s} is,

$$Var(XM_{1s}) = Var(X)Var(M_{1s}) - Cov(X, M_{1s})^2.$$

(9)

For our specific model, Eq. 9 simplifies because X is a perfectly reliable, randomized variable, so it does not covary with the pretest measures of M_{1s} (i.e., $Cov(X, M_{1s}) = 0$), and its variance does not have to be corrected for measurement error. To correct $Var(XM_{1s})$, we use the corrected variance of M_{1s} (which is an entry in $\Sigma_{M_{sc}}$) in Eq. 9 instead of the observed variance of M_{1s}. Similar steps would follow to correct $Var(XY_{1s})$, $Var(XM_{1f})$, and $Var(XY_{1f})$.

---

[1] The general term in the right hand side of the equation is $\left(B_w\Lambda_M\right)^{-1}[\Sigma_{M_f} - B_w'\Theta_{eM}B_w']\left(\Lambda_M'B_w'\right)^{-1}$ but Bartlett scores have the property that $B_w\Lambda = I$, which simplifies the expression. Regression scores do not have this property.

[2] In Eq. 7 and 8 we are summing the residual variances to estimate the error variance of the observed score. By subtracting the residual variance from the observed score variance, what remains is the reliable variance.

Once the variances and covariances of the scores and the interaction terms are corrected, the final step is to fit the mediation model with interactions (Eq. 1 and 2) to the full corrected covariance matrix using any structural equation modeling program. Note that by fitting the model to the full covariance matrix with maximum likelihood, we assume multivariate normality.

Structural models for latent BTIs

Another option to estimate latent BTI effects is to use structural models. Below we apply three approaches to the two-wave mediation model that have been shown to recover parameters well in single mediator models with a latent interaction (Gonzalez & Valente, in press).

Unconstrained product indicator (UPI) approach.

Product indicator methods estimate latent interaction effects by specifying a latent interaction variable (e.g., a latent variable for XM_1), whose indicators are products of the indicators of the variables involved (Kenny & Judd, 1984; Marsh et al., 2012). There are many ways to specify latent interactions with product indicators (e.g., Aytürk et al., 2021; Jaccard & Wan, 1995; Jöreskog & Yang, 1996; Kenny & Judd, 1984; Marsh et al., 2004; Ping, 1996). In this paper, we use the unconstrained product indicator (UPI) approach (**Method 3**; Marsh et al., 2004). For our specific model, indicators of the XM_1 latent variable were the product of each M_1 indicator and X, the mean of latent XM_1 was fixed to $Cov(X, M_1) = 0$ (because X is a randomized treatment), and the variance of latent XM_1 was fixed to $Var(X)Var(M_1)$ per Eq. 9. The latent variable for XY_1 can be similarly specified. We chose the UPI approach because there are no constraints placed on the estimation of the factor loadings and residual variances of the indicators of the latent interaction terms, which differs from other product indicator methods. Note that the UPI approach relies on multivariate normality.

Latent Moderated Structural Equations (LMS)

Latent Moderated Structural Equations (LMS) is an indirect approach to estimate latent interaction effects (**Method 4**; Klein & Moosbrugger, 2000) and is available to M*plus* users (Muthen & Muthen, 2002–2018). LMS approximates the likelihood of the model with latent interactions using a mixture of normal distributions. A mixture of distributions is used because the conditional distribution of the outcome given the predictors is nonnormal due to the nonlinear relations present in the model (i.e., the linear-by-linear interactions; Kelava et al., 2011). We can represent this model by separating Eq. 1 and 2 into linear and nonlinear effects:

$$M_2 = \begin{bmatrix} a & s_m & c_y \end{bmatrix} \begin{bmatrix} X \\ M_1 \\ Y_1 \end{bmatrix} + \begin{bmatrix} X & M_1 & Y_1 \end{bmatrix} \begin{bmatrix} 0 & h_1 & h_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X \\ M_1 \\ Y_1 \end{bmatrix} + e_M$$

(10)

$$Y_2 = \begin{bmatrix} c' & b & c_m & s_y \end{bmatrix} \begin{bmatrix} X \\ M_2 \\ M_1 \\ Y_1 \end{bmatrix} + \begin{bmatrix} X & M_2 & M_1 & Y_1 \end{bmatrix} \begin{bmatrix} 0 & 0 & h_3 & h_4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X \\ M_2 \\ M_1 \\ Y_1 \end{bmatrix} + e_Y \; .$$

(11)

The model is estimated using the EM algorithm (Dempster et al., 1979, Ng & Chan, 2020). Details of the procedure are highly technical, so we refer interested readers to Kelava et al., (2011) and Klein and Moosbrugger (2000). Note that LMS also relies on multivariate normality.

Bayesian mediation

Bayesian approaches have been shown to yield accurate parameter estimates from mediation models (**Method 5**; Enders et al., 2012; Miočević, 2019; Yuan & MacKinnon, 2009) and handle latent interactions (Asparouhov & Muthén, 2021; Miočević et al., 2018). In general, we can sample from posterior distributions of the model parameters in a mediation model with latent variables (with no interactions) using Markov-chain Monte Carlo (MCMC) estimation (Miočević, 2019). An intermediate step in the estimation of the statistical mediation model with latent variables using a Gibbs sampler is the sampling of a latent variable score for each latent variable (Miočević, 2019). As such, we can incorporate latent BTIs in a Bayesian mediation model by multiplying in each iteration the sampled value of the latent M_1 or Y_1 score by X and regress the posttest mediator and outcome on those values.

Present Study

Examining baseline-by-treatment interactions in statistical mediation analysis can provide insight into how and for whom an intervention works (MacKinnon, 2008). However, it is unknown whether the methods that have been shown to recover parameter estimates well in a single mediator model with one latent interaction (Cheung & Lau, 2017; Gonzalez & Valente, in press) will also recover parameters well in a more complex model, such as the two-wave mediator model with two latent interactions and four BTI effects. Continuing, we are not aware of any resource that generally describes how to incorporate latent BTI effects in mediation analysis and that provides syntax to guide researchers on the estimation of these theoretically relevant models to intervention research. Below, we conduct a simulation study to examine the bias, power, type 1 error rate, and interval coverage of the estimates in the two-wave mediation model and the latent BTIs, illustrate the methods with an applied intervention example and provide syntax in the supplement to fit these models.

Method

Simulation Factors

True covariance matrices for the two-wave mediation model with baseline-by-treatment interactions were derived and are shown in the supplement. Datasets were generated in the R statistical environment by varying the factors found in Table 1. Factors that did not vary

in the simulation were the effect size of the c'-path (small), the stabilities of M and Y (at r = . 7), the cross-lag effects (zero effect), and the correlation of M_1 and Y_1 at baseline (r = . 575). Overall, there were 96 conditions, with 500 replications per condition.

All path effect sizes were obtained from Valente & MacKinnon (2017) and approximately match Cohen's f^2 benchmarks. Residual covariance values were specified to be similar to Cohen's effect sizes of small and medium correlation, and the sample sizes studied are achievable in social science applications and have been found to recover parameter estimates well in a single mediator model with one latent interaction (Gonzalez & Valente, in press). Items for M_1, Y_1, M_2, and Y_2 were simulated to be unidimensional and adhere to strict invariance across time. All item intercepts were zero, factor loadings for M_1, Y_1, M_2, and Y_2 were drawn from a uniform random distribution with limits between .5 and .7, and the residual variances were the complement of the factor loadings so that the total variance of each item at baseline was equal to 1. Note that all the indicators for the conditions presented here are multivariate normally distributed, so they meet the distributional assumptions of the methods examined. In the supplement, we show that the stability paths are biased when discrete, nonnormally distributed indicators are treated as continuous. Therefore, we do not recommend using these methods in these scenarios.

General Procedure

Each dataset was analyzed with each of the five methods to estimate the latent BTI effect: summed scores with reliability adjustment, factor scores (Bartlett) with reliability adjustment, unconstrained product indicator (UPI) approach, latent moderated structural equations (LMS), and Bayesian mediation analysis. Note that all longitudinal latent variable models fit to the data were specified to be scalar invariant and correlated residuals were prespecified. For the summed scores and factor score methods, we used the parameter estimates from the longitudinal factor model for M_1 and M_2 and the longitudinal factor model for Y_1 and Y_2 to correct the covariance matrices as discussed above. Then, the statistical mediation model with observed BTIs was fit to both corrected covariance matrices using `lavaan` in R (Rosseel, 2012). Furthermore, the LMS approach (using the XWITH argument for latent interactions, random effect estimation, and numerical integration) and UPI approach (with constraints explained above) were estimated in M*plus* and processed in R using the `MplusAutomation` package (Hallquist & Wiley, 2020). Finally, the Bayesian mediation model was estimated in JAGS (Plummer, 2003) in R using the `R2jags` package (Su & Yajima, 2020).[3] Similar to Gonzalez & Valente (in press), we specified diffuse prior distributions for model parameters to mimic situations with limited prior beliefs about the parameters: factor loadings had normal distributions N(1, 20) truncated at 0 (i.e., the loadings were positive), item intercepts had normal distributions N(0, 20), residual variances had inverse gamma distributions IG(1,1), latent variable means for exogenous variables had a normal distribution N(0,1), and regression coefficients and equation intercepts had

---

[3]We imposed longitudinal scalar invariance constraints by using the same parameter in different parts of the model. For example, a single posterior distribution represented the relation between M_1 and the item m_{11} and M_2 and the item m_{21}. Also, we modeled the correlated residuals like the specific factors in bifactor models: the factor had two indicators, which were the same item across time, and these factors were orthogonal to all other variables.

normal distributions N(0, 1000). Pilot analyses suggested that the chains for each parameter achieved a potential scale reduction (PSR) factor < 1.1 with at most 5,000 iterations. As such, we used three chains which each sample 5,000 MCMC iterations (1,000 of which were burn-in iterations).

Monte Carlo Outcomes

We examined three bias outcomes. Raw bias was computed in each replication by the difference between the estimated parameter and the data-generating value.[4] We also report relative bias (i.e., raw bias divided by the true value of the estimate) for conditions in which the true value of the parameter is nonzero, and standardized bias (i.e., raw bias divided by the standard deviation of the parameter estimates across replications per condition) for conditions in which the true value of the parameter is zero. Bias outcomes were analyzed using a mixed ANOVA as a function of the simulation factors as between-group factors, the latent variable method as a within-group factor, and all possible interactions. Conditions with relative bias > .10 (Kaplan, 1988) or standardized bias > .40 (Collins et al., 2001), and predictors with partial-η^2 > .005 were further discussed. We also estimated 95% nonparametric percentile bootstrap confidence intervals (with 500 bootstrap draws) for model parameters and the mediated effect across all methods (MacKinnon et al., 2004) except Bayesian mediation, in which equal-tail credible intervals based on the posterior distributions were used (Yuan & MacKinnon, 2009). Power was defined as the proportion of replicates in which the interval did not contain zero for parameters with nonzero true values, and type 1 error was the proportion of replicates in which the interval did not contain zero for parameters with zero true values. Note that power was estimated for the a-path, b-path, c-path, stability estimates, and the BTI effects, while type 1 error was only estimated for the BTI effects and cross-lag effects. Finally, interval coverage was defined as the proportion of replicates where the derived interval contained the true value of the parameter. Power, type 1 error rates, and interval coverage were analyzed with regression trees (Gonzalez et al., 2018), where splits had to yield an increase of at least $R^2 = .005$ on the outcome to be interpreted.

Results

Parameter Bias and Interval Coverage

Table 2 shows the parameter bias for the mediated effect and the BTI coefficients. In the ANOVAs predicting parameter bias (i.e., raw, relative, and standardized bias) from all simulation factors and methods used to estimate the latent interaction effect, there were three parameters with factors that had effect sizes with $\eta^2 > .005$. There were main effects for the method used on the raw bias of the a-path $\left(\eta^2 = 0.007\right)$, c-path $\left(\eta^2 = 0.007\right)$, cross-lag effect of Y $\left(\eta^2 = 0.011\right)$, and on the raw bias $\left(\eta^2 = 0.041\right)$ and relative bias $\left(\eta^2 = 0.041\right)$ of the stability of Y. Also, there were significant interactions between the method and sample size and the method and the number of items on the raw bias ($\eta^2 = 0.014$ and 0.030, respectively)

---

[4]Rescaled true values for the corrected summed score model were derived to serve as reference values for raw bias. Summed scores are on a different metric than the latent variable models, similar to Georgeson et al., (2021).

and relative bias ($\eta^2 = 0.014$ and 0.030, respectively) of the stability of M, and a significant interaction between the method and sample size for the standardized bias on the cross-lag effect of Y $\left(\eta^2 = 0.007\right)$. However, across methods, all the relative bias estimates for nonzero parameters were below .05, and all the standardized bias estimates for zero parameters were below .40. Furthermore, Table 2 shows the interval coverage across methods. Regression trees predicting coverage from the simulation factors and the BTI estimation method did not yield any splits, so there were no important differences in interval coverage across simulation factors. All interval coverage estimates are within the range of .925 and .975.

Statistical Power and Type 1 error rate

Regression trees for the analysis of statistical power and type 1 error by the simulation factor are shown in the supplement. For statistical power, the effect size of the parameter and sample size were significant predictors, as expected. Furthermore, the h1 coefficient (e.g., the relation between the XM_1 interaction and M_2) had higher power in conditions with six items than with four items at N = 500. Recall that all BTI effects had a small effect size. Power was .310 and .430 for conditions with four and six items in the mediator, respectively. We presume that difference in power might be because more items yield more stable latent variables. Also, the c-path had higher power in conditions when the a-path had a small effect size (power $= .750$) than in conditions with a medium effect size (power $= .650$). Although Table 3 might suggest that there are some power differences across methods, the method to estimate the latent interaction effect was not an important predictor in the trees. On the other hand, results suggest that there were no significant predictors for type 1 error rate for the BTI estimates and cross-lag effects. Table 4 shows the type 1 error rates across methods, and all the estimates were within the bounds of .025 and .075.

Summary

Our results suggest that, as long as the indicators are multivariate normally distributed, summed scores with reliability adjustment, factor scores with reliability adjustment, UPI, LMS, and Bayesian mediation provide unbiased parameter estimates in the mediation model, the 95% confidence/credible interval coverages perform appropriately, and the type 1 error rates were close to .05. Furthermore, methods did not differ significantly on parameter bias, power, type 1 error rates, and interval coverage. As such, similar to the findings in the single mediator model with one latent interaction, the studied methods (which include our proposed extension to the scoring approaches to adjust for measurement error and the effect of correlated residuals simultaneously) recover the parameter estimates well.

Illustration

Data were obtained from The Job Search Intervention Study (JOBS II), which is a randomized field intervention study that, among other aims, examined if a job training intervention for unemployed workers improved symptoms of depression (outcome) by enhancing job search self-efficacy (mediator), controlling for baseline covariates (Vinokur et al., 1995; Vinokur & Schul, 1997). The original sample size was N = 1,801, but to simplify the analyses, we used only participants who had complete item-level data for the mediator and outcome at both timepoints. Therefore, the sample size used for this illustration was

N = 1126, with N = 347 in the control group and N = 779 in the treatment group. Depressive symptoms were measured by the eleven five-category items on the Hopkins Symptoms Checklist (coefficient $\alpha = .807$ at pretest and $\alpha = .884$ at posttest), which asked about the frequency of several symptoms, such as loneliness and hopelessness, experienced by the respondent in the past two weeks. Higher scores represented more depressive symptoms. The job search self-efficacy measure had six five-category items (coefficient $\alpha = .868$ at pretest and $\alpha = .886$ at posttest), which asked how confident respondents were about several aspects of the job process, such as completing a job application and a resume. Higher scores represented higher job search self-efficacy. Preliminary analyses suggest that the mediator and outcome measures are largely unidimensional and meet the longitudinal scalar invariance assumptions (see supplement). Here, we test a statistical mediation model with latent baseline-by-treatment interactions to investigate if the effect of the training on job search self-efficacy and depressive symptoms depends on the baseline levels of those constructs. The methods examined were summed scores with reliability adjustment, factor scores with reliability adjustment, UPI, LMS, and Bayesian mediation, and used either 95% percentile bootstrap confidence intervals or 95% equal-tail credible intervals for statistical inference. The specification of the models, including the prior distributions of parameters in Bayesian mediation, closely follow the specification described for the simulation.

Results

Table 5 shows the parameter estimates of the mediation model from Eq. 1 and 2. The estimates were similar across models, so we will only interpret the estimates for the factor scores with reliability adjustment. The treatment significantly increased job search self-efficacy ($a = .267$, 95%CI = [.151, .359]), and job search self-efficacy was significantly associated with lower depressive symptoms ($b = -.224$, 95%CI = [-.390, -.163]). There was also a significant baseline-by-treatment interaction on the mediator at posttest (XM_1; h1 = $-.144$, 95%CI = [-.250, -.027]), which indicated that the relation between the program and job search self-efficacy at posttest was stronger for individuals who had low job search self-efficacy at pretest. There was a significant mediated effect at the mean of M_1 ($ab = -.059$, 95%CI = [-.118, -.029]), and a significant mediated effect for individuals 1 SD below the mean of M_1 ($ab = -.104$, 95%CI = [-.202, -.103]), but the interaction was not significant for individuals 1 SD above the mean of M_1 ($ab = -.016$, 95%CI = [-.059, .016]). Given these findings, future implementations of the intervention might consider screening for job search self-efficacy before enrolling individuals in the intervention because individuals with low job search self-efficacy might benefit more from the program.

Discussion

Estimating baseline-by-treatment interactions can reveal how and for whom an intervention works, which is valuable information for the dissemination and implementation of interventions. In this paper, we discussed five different methods (e.g., summed scores with reliability adjustment, factor scores with reliability adjustment, unconstrained product indicator approach, latent moderated structural equations, and Bayesian mediation) for estimating the statistical mediation model with baseline-by-treatment interactions when the

mediator and outcome are latent variables. We also evaluated the estimation properties of the approaches and illustrated the methods using an applied example and presented code in the supplement to facilitate testing for latent baseline-by treatment interactions. Our simulation results suggest that the five approaches examined led to unbiased effects, similar power, interval coverage of approximately .95, and type 1 error rates close to .05. As such, each of these methods estimated the mediation effects and the latent BTI effects accurately. However, across methods, the estimates of power to detect small effect sizes were well below .80 even for a condition with N = 500, which might be a limiting factor in applied research. Larger sample sizes might be needed to detect small effects using this model.

Given that all methods provided accurate results, additional simulations may reveal situations in which they provide different answers. Determining which latent interaction approach to use with any given model remains an open question (Kelava & Brandt, 2022). As such, researchers estimating latent BTI effects in mediation models could follow current recommendations from the literature. In a recent review of work in latent interactions, Kelava and Brandt (2022) recommend using the LMS approach in situations in which there are normally-distributed latent variables and indicators. Kelava and Brandt (2022) mention that LMS is more efficient than the UPI approach (there is a loss of degrees of freedom when the estimation of the latent interaction term is unconstrained) and the Bayesian mediation approach (where efficiency depends on the specified prior), and that more research is needed to understand the performance of the scoring approaches. Other considerations include amount of background, analytical flexibility, computational time, and programming needed for their implementation. For example, the specification of the LMS procedure is relatively simple and took about 1–3 minutes to estimate a single model, and the procedure is available in M*plus*. If access to M*plus* is a limitation, then researchers could use any of the other methods presented – Bayesian mediation can be used in any program that does MCMC estimation, while the others can be estimated in any structural equation modeling program that takes a covariance matrix as an input. However, estimating the Bayesian mediation model would require researchers to be familiar with Bayesian inference and specifications of prior distributions, which might pose a challenge, along with long computation times. For the model in the illustration, the estimation and parallelization per MCMC chain took around 50 minutes in JAGS using a 13in MacBook Pro 2020 with 32GB of RAM. On the other hand, the coding to estimate the model using the UPI approach or summed scores and factor scores with reliability adjustment could be cumbersome, but the estimation time is almost immediate. As such, we advise researchers to review the code we provide in the appendix to gain familiarity with each of these approaches.

Extensions to the Model

In applied settings, researchers often use extensions of the two-wave model examined here to test specific theories. Depending on the model and situation, these extensions might or might not be easily accommodated by the procedures examined. These extensions require more simulation work to determine their performance, but we list how the methods examined could handle these extensions. For example, it may be of interest to include other covariates in Eq. 1 and 2 to increase statistical power (e.g., Gonzales et al., 2012). In cases in which the covariates are observed and perfectly reliable, the extensions are simpler. If

using summed scores and factor scores with reliability adjustment, the covariates would be included in the covariance matrix that is analyzed and additional adjustments to the covariance matrix would not be necessary. For the UPI, LMS, and Bayesian methods, the observed covariate could be specified and included in the corresponding equations. In cases in which the covariates are latent or have measurement error, adding a covariate adds another layer of complexity. The methods with summed scores and factor scores with reliability adjustment would need to score the covariate, correct its variance for measurement error, and include it in the analyzed covariance matrix. For the UPI, LMS, and Bayesian mediation, a latent covariate could be specified and included in the corresponding equations.

Furthermore, the methods examined could be extended to situations in which M and Y are measured at baseline and at two additional timepoints (i.e., three-wave experimental designs). A third wave adds a layer of complexity to the two-wave model because M3 and Y3 are two additional latent variables to estimate. Perhaps in this design, the mediated effect of most interest is the X → M2 → Y3 relation, where each path can be moderated by baseline levels of M and Y. The methods with summed scores and factor scores with reliability adjustments would correct the 3×3 variance-covariance submatrices of M and Y for measurement error and correlated residuals. If interested, researchers can take the steps to correct the XM1 and XY1 effects in the former Y2 equation and correct the XM2 and XY2 effects in the Y3 equation. For the UPI, LMS, and Bayesian methods, latent M3 and Y3 could be specified and included in the models.

Finally, the methods examined could be extended to situations in which X is a continuous variable (e.g., an intervention fidelity indicator). Although the procedures would stay the same when X is perfectly reliable, the procedures would differ when X has measurement error or is represented by a latent variable. The methods with summed scores and factor scores with reliability adjustments would need to correct the variance of X for measurement error, and consequently the variances of XM1 and XY1, in the analyzed covariance matrix. For the UPI, LMS, and Bayesian methods, a (single-indicator or multiple-indicator) latent variable could be specified for X and latent interactions for XM1 and XY1 and then the researcher would proceed as described. Note that we assume multivariate normality for all of these effects. Beyond the modeling complexity, it is important to note that the nonrandomized nature of X opens the door to unmeasured confounders in the X → M2 and X → Y2 relations.

Limitations and Future Directions

An important limitation of this study is that the results we presented rely on the correct specification of the models. A future direction is to investigate how longitudinal noninvariance (Georgeson et al., 2021), unmodeled multidimensionality (Gonzalez & MacKinnon, 2018), distribution of the latent variables (Cham et al., 2012), or structural misspecifications (Kelava & Brandt, 2022) affect the estimation of the statistical mediation model with latent BTIs. Specifically, it would be important to study whether ignoring these aspects leads to spurious observed baseline-by-treatment interactions. Additionally, we could have studied how sensitive our results are to different specifications of the models. For example, one could study how sensitive our results are to prior distributions used in the

Bayesian mediation model. In our study, diffuse priors were able to recover the parameters well, but this might not always be the case. Furthermore, although previous research suggests that varying stability does not affect the accuracy of the mediated effect estimate in two-wave mediation models without interactions (Valente & MacKinnon, 2017), one could examine if these findings also hold in two-wave mediation models with latent BTIs. Also, it would be interesting for future studies to increase the number of replications per simulation condition to obtain more precise estimates of our Monte Carlo outcomes. Finally, there are other methods to handle latent BTIs that we did not study, such as two-stage least squares methods (Bollen & Paxton, 1998) or others discussed by Kelava and Brandt (2022). Overall, we encourage researchers to estimate statistical mediation models with latent baseline-by-treatment interactions using any of the methods discussed in this paper for accurate estimation of the effects and to help with the implementation and dissemination of interventions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

References

Aiken LS, & West SG (1991). Multiple regression: Testing and interpreting interactions. Newbury Park, CA: Sage.

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.

Aytürk E, Cham H, Jennings PA, & Brown JL (2021). Exploring the performance of latent moderated structural equations approach for ordered-categorical items. Structural Equation Modeling: A Multidisciplinary Journal, 28, 410–422.

Bartlett MS (1937). The statistical conception of mental factors. British Journal of Psychology, 28, 97–104.

Baron RM, & Kenny DA (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology, 51, 1173–1182. [PubMed: 3806354]

Bohrnstedt GW, & Marwell G (1978). The reliability of products of two random variables. Sociological Methodology, 9, 254–273.

Bollen KA, & Paxton P (1998). Interactions of latent variables in structural equation models. Structural Equation Modeling: A Multidisciplinary Journal, 5, 267–293.

Busemeyer JR, & Jones LE (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. Psychological Bulletin, 93, 549–562.

Cham H, West SG, Ma Y, & Aiken LS (2012). Estimating latent variable interactions with nonnormal observed data: A comparison of four approaches. Multivariate Behavioral Research, 47(6), 840–876. [PubMed: 23457417]

Chaudoir SR, Behari K, Williams SL, & Pachankis JE (2021). Why do brief online writing interventions improve health? Examining mediators of expressive writing and self-affirmation intervention efficacy among sexual minority emerging adults. Psychology of Sexual Orientation and Gender Diversity. Advance online publication. 10.1037/sgd0000507

Cheung GW, & Lau RS (2017). Accuracy of parameter estimates and confidence intervals in moderated mediation models: A comparison of regression and latent moderated structural equations. Organizational Research Methods, 20(4), 746–769.

Cole DA, & Preacher KJ (2014). Manifest variable path analysis: potentially serious and misleading consequences due to uncorrected measurement error. Psychological Methods, 19, 300–315. [PubMed: 24079927]

Croon M (2002). Using predicted latent scores in general latent structure models. In Marcoulides G & Moustaki I (Eds.), Latent variable and latent structure modeling (pp. 195–223). Mahwah, NJ: Erlbaum.

Collins LM, Schafer JL, & Kam C-M (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychological Methods, 6, 330–351. [PubMed: 11778676]

Cox K, & Kelcey B (2021). Croon's bias-corrected estimation of latent interactions. Structural Equation Modeling: A Multidisciplinary Journal, 1–12.

Cox MG, Kisbu-Sakarya Y, Mio evi M, & MacKinnon DP (2013). Sensitivity plots for confounder bias in the single mediator model. Evaluation Review, 37, 405–431. [PubMed: 24681690]

Dempster AP, Laird NM, & Rubin DB (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Methodological), 39, 1–22.

Enders CK, Fairchild AJ, & MacKinnon DP (2013). A Bayesian approach for estimating mediation effects with missing data. Multivariate Behavioral Research, 48, 340–369. [PubMed: 24039298]

Fritz MS, Kenny DA, & MacKinnon DP (2016). The combined effects of measurement error and omitting confounders in the single-mediator model. Multivariate Behavioral Research, 51, 681–697. [PubMed: 27739903]

Fritz MS, MacKinnon DP, Williams J, Goldberg L, Moe EL, & Elliot DL (2005). Analysis of baseline by treatment interactions in a drug prevention and health promotion program for high school male athletes. Addictive Behaviors, 30, 1001–1005. [PubMed: 15893095]

Georgeson AR, Valente MJ, & Gonzalez O (2021). Evaluating response shift in statistical mediation analysis. Advances in Methods and Practices in Psychological Science, 4, 1–13.

Gonzales NA, Dumka LE, Millsap RE, Gottschall A, McClain DB, Wong JJ, ... & Kim SY (2012). Randomized trial of a broad preventive intervention for Mexican American adolescents. Journal of Consulting and Clinical psychology, 80(1), 1–16. [PubMed: 22103956]

Gonzalez O & MacKinnon DP (2018). A bifactor approach to model multifaceted constructs in statistical mediation analysis. Educational and Psychological Measurement, 78, 5–31. [PubMed: 29335655]

Gonzalez O, & MacKinnon DP (2021). The measurement of the mediator and its influence on statistical mediation conclusions. Psychological Methods, 26, 1–17. [PubMed: 32175754]

Gonzalez O, O'Rourke HP, Wurpts IC, & Grimm KJ (2018). Analyzing Monte Carlo simulation studies with Classification and Regression Trees. Structural Equation Modeling: A Multidisciplinary Journal, 25, 403–413.

Gonzalez O & Valente MJ (in press). Accommodating a latent XM interaction in statistical mediation analysis. Multivariate Behavioral Research. DOI: 10.1080/00273171.2022.2119928.

Hallquist MN, & Wiley JF (2018). MplusAutomation: an R package for facilitating largescale latent variable analyses in Mplus. Structural Equation Modeling: A Multidisciplinary Journal, 25, 621–638. [PubMed: 30083048]

Hayes AF (2015). An index and test of linear moderated mediation. Multivariate Behavioral Research, 50(1), 1–22. [PubMed: 26609740]

Hayes T, & Usami S (2020). Factor score regression in the presence of correlated unique factors. Educational and Psychological Measurement, 80, 5–40. [PubMed: 31933491]

Howe GW (2019). Using baseline target moderation to guide decisions on adapting prevention programs. Development and psychopathology, 31, 1777–1788. [PubMed: 31496463]

Hoyle RH & Kenny DA (1999). Sample size, reliability, and tests of statistical mediation. In Hoyle RH (Ed.), Statistical strategies for small sample research (pp. 195–222). Thousand Oaks, CA: Sage.

Hsiao Y-Y, Kwok OM, & Lai MH (2018). Evaluation of two methods for modeling measurement errors when testing interaction effects with observed composite scores. Educational and Psychological Measurement, 78, 181–202. [PubMed: 29795952]

Imai K, Keele L, & Tingley D (2010). A general approach to causal mediation analysis. Psychological Methods, 15, 309–334. [PubMed: 20954780]

Jaccard J, & Wan CK (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. Psychological Bulletin, 117, 348–357.

Jensen MR, Wong JJ, Gonzales NA, Dumka LE, Millsap R, & Coxe S (2014). Long-term effects of a universal family intervention: Mediation through parent-adolescent conflict. Journal of Clinical Child & Adolescent Psychology, 43, 415–427. [PubMed: 24730357]

Jöreskog KG, & Yang F (1996). Nonlinear structural equation models: The Kenny–Judd model with interaction effects. In Marcoulides GA & Schumacker RE(Eds.), Advanced structural equation modeling (pp. 57–88). Hillsdale, NJ: Erlbaum.

Kaplan D (1988). The impact of specification error on the estimation, testing, and improvement of structural equation models. Multivariate Behavioral Research, 23, 69–86. [PubMed: 26782258]

Kelava A, & Brandt H, (2022) Latent interaction effects. In Hoyle R (Ed.) Handbook of Structural Equation Modeling (Vol. 2, pp. 427–446). The Guilford Press.

Kelava A, Werner CS, Schermelleh-Engel K, Moosbrugger H, Zapf D, Ma Y, Cham H, Aiken LS, & West SG (2011). Advanced nonlinear latent variable modeling: Distribution analytic LMS and QML estimators of interaction and quadratic effects. Structural equation Modeling: A Multidisciplinary Journal, 18, 465–491.

Kenny DA, & Judd CM (1984). Estimating the nonlinear and interactive effects of latent variables. Psychological Bulletin, 96, 201–210.

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

Kunze AE, Lancee J, Morina N, Kindt M, & Arntz A (2019). Mediators of change in imagery rescripting and imaginal exposure for nightmares: evidence from a randomized wait-list controlled trial. Behavior Therapy, 50(5), 978–993. [PubMed: 31422852]

Lawley DN, & Maxwell AE (1971). Factor analysis as a statistical method. New York: American Elsevier.

le Cessie S, Debeij J, Rosendaal FR, Cannegieter SC, & Vandenbrouckea JP (2012). Quantification of bias in direct effects estimates due to different types of measurement error in the mediator. Epidemiology, 551–560. [PubMed: 22526092]

Levy R, & Mislevy RJ (2016). Bayesian psychometric modeling. CRC Press.

Liu X, & Wang L (2021). The impact of measurement error and omitting confounders on statistical inference of mediation effects and tools for sensitivity analysis. Psychological Methods, 26, 327–342. [PubMed: 32718152]

MacKinnon DP (2008). Introduction to statistical mediation analysis, New York, NY: Lawrence Erlbaum.

MacKinnon DP, Goldberg L, Clarke GN, Elliot DL, Cheong J, Lapin A, ... & Krull, J. L. (2001). Mediating mechanisms in a program to reduce intentions to use anabolic steroids and improve exercise self-efficacy and dietary behavior. Prevention Science, 2, 15–28. [PubMed: 11519372]

MacKinnon DP, Lockwood CM, & Williams J (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. Multivariate Behavioral Research, 39, 99–128. [PubMed: 20157642]

MacKinnon DP, Valente MJ, & Gonzalez O (2020). The correspondence between causal and traditional mediation analysis: The link is the mediator by treatment interaction. Prevention Science, 21, 147–157. [PubMed: 31833021]

Marsh HW, Wen Z, & Hau K-T (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. Psychological Methods, 9, 275–300. [PubMed: 15355150]

Marsh HW, Wen Z, Nagengast B, & Hau K-T (2012). Structural equation models of latent interaction. In Hoyle RH (Ed.), Handbook of structural equation modeling (pp. 436–458). The Guilford Press.

Mayer A, Thoemmes F, Rose N, Steyer R, & West SG (2014). Theory and anlysis of total, direct, and indirect causal effects. Multivariate Behavioral Research, 49(5), 425–442. [PubMed: 26732357]

McClelland GH, & Judd CM (1993). Statistical difficulties of detecting interactions and moderator effects. Psychological Bulletin, 114, 376–390. [PubMed: 8416037]

Miočević M (2019). A tutorial in Bayesian mediation analysis with latent variables. Methodology, 15, 137–146.

Miočević M, Gonzalez O, Valente MJ, & MacKinnon DP (2018). A tutorial on Bayesian potential outcomes mediation analysis. Structural Equation Modeling: A Multidisciplinary Journal, 25, 121–136. [PubMed: 29910595]

Montoya AK (2019). Moderation analysis in two-instance repeated measures designs: Probing methods and multiple moderator models. Behavior Research Methods, 51, 61–82. [PubMed: 30306409]

Morgan-Lopez AA, & MacKinnon DP (2006). Demonstration and evaluation of a method for assessing mediated moderation. Behavior Research Methods, 38(1), 77–87. [PubMed: 16817516]

Muthén LK and Muthén BO (1998–2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

Ng JC, & Chan W (2020). Latent moderation analysis: a factor score approach. Structural Equation Modeling: A Multidisciplinary Journal, 27, 629–648.

Pearl J (2001). Direct and indirect effects. In Breese J & Koller D (Eds.), Proceedings of the 17th Conference on Uncertainty in Artificial Intelligence (pp. 411–420). San Francisco, CA: Morgan Kaufmann.

Perrino T, Pantin H, Prado G, Huang S, Brincks A, Howe G, … & Brown CH. (2014). Preventing internalizing symptoms among Hispanic adolescents: A synthesis across Familias Unidas trials. Prevention Science, 15, 917–928. [PubMed: 24390684]

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

Plummer M (2003, March). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In Proceedings of the 3rd international workshop on distributed statistical computing (Vol. 124, No. 125.10, pp. 1–10).

Rosseel Y (2012). Lavaan: An R package for structural equation modeling and more. Journal of Statistical Software, 48, 1–36.

Skrondal A, & Laake P (2001). Regression among factor scores. Psychometrika, 66, 563–575.

Smith JD, Knoble NB, Zerr AA, Dishion TJ, & Stormshak EA (2014). Family check-up effects across diverse ethnic groups: Reducing early-adolescence antisocial behavior by reducing family conflict. Journal of Clinical Child & Adolescent Psychology, 43, 400–414. [PubMed: 24731120]

Su Y-S, & Yajima M (2020). R2jags: Using R to Run 'JAGS'. R package version 0.6–1. https://CRAN.R-project.org/package=R2jags.

te Brinke LW, Menting AT, Schuiringa HD, Deković M, Weisz JR, & De Castro BO (2021). Emotion regulation training as a treatment element for externalizing problems in adolescence: A randomized controlled micro-trial. Behaviour Research and Therapy, 143, 103889. [PubMed: 34111699]

Tofighi D, & MacKinnon DP (2011). RMediation: An R package for mediation analysis confidence intervals. Behavior Research Methods, 43, 692–700. [PubMed: 21487904]

Valente MJ, & MacKinnon DP (2017). Comparing models of change to estimate the mediated effect in the pretest–posttest control group design. Structural Equation Modeling: A Multidisciplinary Journal, 24, 428–450. [PubMed: 28845097]

Valente MJ, MacKinnon DP, & Mazza GL (2020). A viable alternative when propensity scores fail: Evaluation of inverse propensity weighting and sequential G-estimation in a two-wave mediation model. Multivariate Behavioral Research, 55, 165–187. [PubMed: 31220937]

Valeri L & VanderWeele TJ (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. Psychological Methods, 18, 137–150. [PubMed: 23379553]

VanderWeele T (2015). Explanation in causal inference: Methods for mediation and interaction. New York, NY: Oxford University Press.

VanderWeele TJ, Valeri L, & Ogburn EL (2012). The role of measurement error and misclassification in mediation analysis. Epidemiology, 23, 561–564. [PubMed: 22659547]

Vinokur AD, Price RH, & Schul Y (1995). Impact of the JOBS intervention on unemployed workers varying in risk for depression. American Journal of Community Psychology, 23, 39–74. [PubMed: 7572826]

Vinokur AD, & Schul Y (1997). Mastery and inoculation against setbacks as active ingredients in the JOBS intervention for the unemployed. Journal of Consulting and Clinical Psychology, 65, 867–877. [PubMed: 9337505]

Wall MM, & Amemiya Y (2003). A method of moments technique for fitting interaction effects in structural equation models. British Journal of Mathematical and Statistical Psychology, 56(1), 47–63. [PubMed: 12803821]

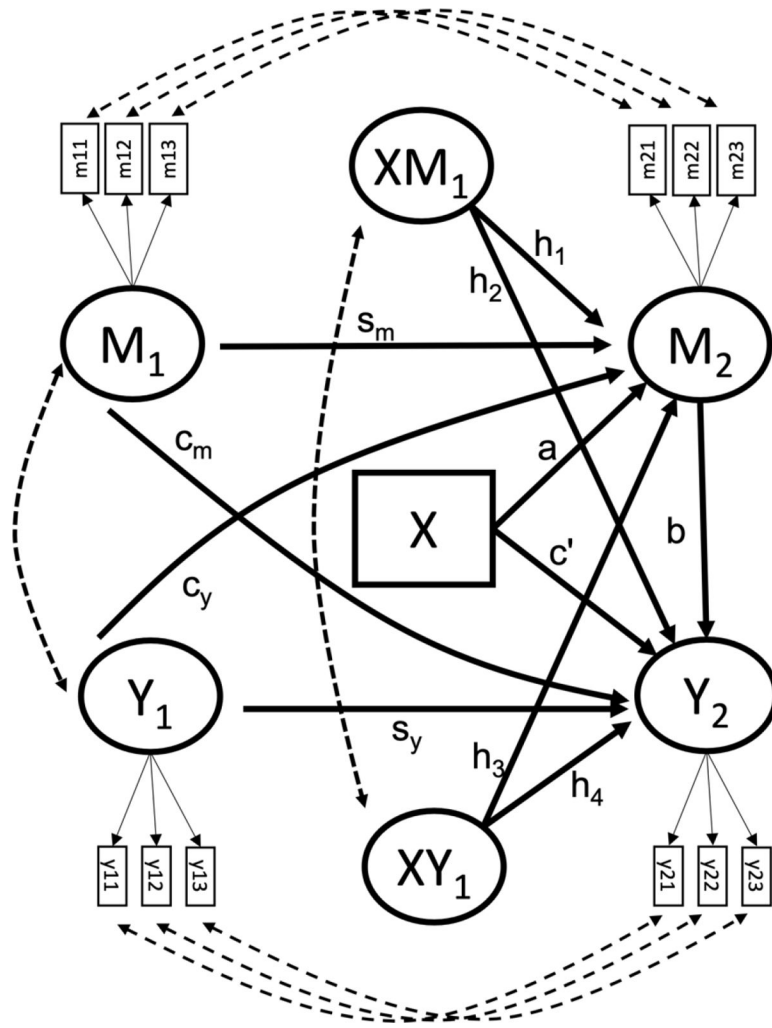Yuan Y, & MacKinnon DP (2009). Bayesian mediation analysis. Psychological Methods, 14, 301–322. [PubMed: 19968395]

Figure 1.
Statistical mediation model with latent baseline-by-treatment interactions and correlated residuals.

Table 1.

Simulation factors

Factor (Levels)	Values
sample size (2)	250, 500
a-path effect size (2)	0.285 (S), 0.845 (M)
b-path effect size (2)	0.145 (S), 0.435 (M)
h-path effect size (3)	[h1-h4 = 0] [h1-h4 = 0.30 (S)] [h1 & h4 = 0.30 (S), h2 & h3 = 0]
# of items (2)	[4 for M and 8 for Y] [6 for M and 6 for Y]
Residual covariance (2)	0.10, 0.30

Note: S is small effect size; M is medium effect size

Table 2.

Bias estimates and interval coverage of the parameters from the mediation model with latent BTIs

	a-path	b-path	c'-path	sm-path	sy-path	ab	cy-path	h1	h2	cm-path	h3	h4
			Relative Bias							*Standardized Bias*		
Summed	−0.003	0.010	−0.005	0.018	0.009	0.010	−0.021	0.001	0.001	−0.004	0.000	0.000
Factor	−0.001	0.005	0.002	0.015	0.009	0.003	−0.003	−0.010	0.002	−0.003	0.001	−0.007
UPI	0.011	−0.001	0.013	0.008	0.004	0.011	0.020	0.005	0.006	0.011	0.002	0.008
LMS	0.009	−0.001	0.011	0.008	0.003	0.009	0.015	0.002	0.001	0.016	−0.002	0.002
Bayes	0.011	−0.002	0.006	0.025	0.012	0.009	−0.027	0.008	0.007	−0.009	−0.002	0.005
			Raw Bias									
Summed	−0.002	0.005	−0.007	0.017	0.007	0.005	−0.007	0.005	−0.001	−0.011	−0.001	0.000
Factor	0.000	0.001	0.001	0.015	0.008	0.000	−0.008	0.001	−0.001	−0.006	0.000	−0.002
UPI	0.007	0.000	0.004	0.008	0.004	0.001	−0.004	0.003	0.000	−0.001	0.000	0.002
LMS	0.006	−0.001	0.003	0.008	0.003	0.001	−0.004	0.002	−0.001	0.000	−0.001	0.001
Bayes	0.007	−0.001	0.002	0.025	0.010	0.001	−0.011	0.003	0.001	−0.004	−0.001	0.001
							Interval Coverage					
Summed	0.951	0.948	0.953	0.943	0.948	0.952	0.948	0.955	0.951	0.947	0.952	0.953
Factor	0.948	0.950	0.952	0.946	0.949	0.950	0.950	0.957	0.953	0.950	0.953	0.953
UPI	0.945	0.948	0.950	0.945	0.946	0.948	0.947	0.947	0.947	0.945	0.949	0.948
LMS	0.946	0.948	0.950	0.949	0.949	0.948	0.949	0.949	0.949	0.948	0.949	0.950
Bayes	0.954	0.952	0.957	0.946	0.951	0.954	0.949	0.949	0.949	0.952	0.951	0.952

Table 3.

Statistical power of the mediation model parameters across methods.

	a-path (s)	a-path (m)	b-path (s)	b-path (m)	c'-path (s)	h1 (s)	h2 (s)	h3 (s)	h4 (s)	ab (s,s)	ab (m,s)	ab (s,m)	ab (m,m)
						N = 250							
Summed	0.381	0.994	0.294	0.938	0.360	0.164	0.186	0.173	0.185	0.096	0.283	0.346	0.931
Factor	0.405	0.997	0.305	0.945	0.391	0.172	0.201	0.192	0.197	0.106	0.293	0.373	0.942
UPI	0.446	0.999	0.320	0.971	0.436	0.212	0.247	0.229	0.225	0.129	0.307	0.424	0.971
LMS	0.443	0.999	0.323	0.972	0.437	0.223	0.247	0.231	0.225	0.128	0.312	0.422	0.973
Bayes	0.393	0.999	0.317	0.977	0.389	0.233	0.254	0.240	0.236	0.108	0.309	0.373	0.976
						N = 500							
Summed	0.683	1.000	0.567	0.999	0.677	0.316	0.400	0.342	0.378	0.355	0.564	0.679	0.999
Factor	0.692	1.000	0.577	0.999	0.693	0.329	0.419	0.363	0.396	0.370	0.578	0.688	0.999
UPI	0.727	1.000	0.585	0.999	0.716	0.383	0.453	0.405	0.416	0.397	0.588	0.724	1.000
LMS	0.727	1.000	0.591	1.000	0.716	0.408	0.461	0.415	0.435	0.403	0.591	0.726	1.000
Bayes	0.696	1.000	0.584	1.000	0.690	0.414	0.464	0.424	0.438	0.378	0.583	0.693	1.000

Note: Power estimates for the stability of M and Y are not shown because the power was almost 1 across conditions. (s) is for small effect size and (m) is for medium effect size.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4.

Type 1 error for parameters from the mediation model across methods

	cy-path	cm-path	h1	h2	h3	h4
			N = 250			
Summed	0.051	0.051	0.046	0.046	0.047	0.043
Factor	0.047	0.047	0.043	0.041	0.045	0.046
UPI	0.051	0.051	0.054	0.051	0.054	0.052
LMS	0.048	0.048	0.051	0.049	0.050	0.049
Bayes	0.050	0.050	0.052	0.050	0.048	0.045
			N = 500			
Summed	0.054	0.054	0.049	0.055	0.056	0.052
Factor	0.053	0.053	0.047	0.055	0.055	0.052
UPI	0.055	0.055	0.057	0.060	0.056	0.056
LMS	0.054	0.054	0.057	0.055	0.055	0.052
Bayes	0.053	0.053	0.058	0.054	0.050	0.049

Table 5.

Parameter estimates with 95% bootstrap confidence intervals [lower limit, upper limits] from a statistical mediation model with latent BTIs in the JOBS II data.

Par	Summed(c)	Factor(c)	LMS	UPI	Bayes
a-path	**1.162 [.633, 1.600]**	**0.267 [.151, .359]**	**0.272 [.135, .416]**	**0.265 [.160, .371]**	**0.272 [.125, .424]**
b-path	**−0.281 [−.426, −.163]**	**−0.224 [−.390, −.163]**	**−0.202 [−.293, −.118]**	**−0.247 [−.378, −.131]**	**−0.203 [−.282, −.119]**
c-path	−0.463 [−1.370, .296]	−0.063 [−.216, .073]	−0.083 [−.252, .076]	−0.070 [−.201, .074]	−0.083 [−.231, .067]
sm-path	**0.518 [.433, .538]**	**0.518 [.413, .529]**	**0.513 [.462, .565]**	**0.516 [.456, .572]**	**0.514 [.452, .578]**
sy-path	**0.619 [.454, .603]**	**0.535 [.436, .581]**	**0.536 [.461, .610]**	**0.534 [.453, .621]**	**0.536 [.462, .609]**
cm-path	0.035 [−.108, .127]	0.009 [−.100, .097]	−0.001 [−.085, .084]	0.008 [−.109, .115]	−0.003 [−.084, .077]
cy-path	**−0.058 [−.106, −.016]**	**−0.065 [−.119, −.017]**	**−0.073 [−.131, −.010]**	**−0.060 [−.125, −.010]**	**−0.073 [−.146, −.002]**
h1-path	**−0.144 [−.250, −.027]**	**−0.197 [−.291, −.074]**	**−0.174 [−.282, −.065]**	**−0.199 [−.322, −.067]**	**−0.178 [−.323, −.033]**
h2-path	−0.020 [−.115, .096]	−0.063 [−.155, .074]	−0.030 [−.173, .109]	−0.050 [−.170, .070]	−0.029 [−.182, .120]
h3-path	−0.023 [−.195, .190]	0.012 [−.141, .168]	0.039 [−.114, .226]	0.036 [−.132, .191]	0.042 [−.110, .198]
h4-path	−0.074 [−.265, .089]	−0.032 [−.216, .126]	−0.024 [−.190, .137]	−0.030 [−.203, .154]	−0.023 [−.182, .130]
ab	**−0.327 [−.574, −.145]**	**−0.059 [−.118, −.029]**	**−0.055 [−.097, −.023]**	**−0.065 [−.112, −.029]**	**−0.055 [−.095, −.022]**

Note: in bold are estimates that are significant at p< .05. (c) stands for corrected. Also, summed score estimates are in a different metric than the estimates of the other approaches.