Psychological Methods
2016, Vol. 21, No. 2, 189–205

1082-989X/16/$12.00 http://dx.doi.org/10.1037/met0000052

Multilevel Structural Equation Models for Assessing Moderation Within and Across Levels of Analysis

Kristopher J. Preacher
Vanderbilt University

Zhen Zhang
Arizona State University

Michael J. Zyphur
University of Melbourne

Social scientists are increasingly interested in multilevel hypotheses, data, and statistical models as well as moderation or interactions among predictors. The result is a focus on hypotheses and tests of multilevel moderation within and across levels of analysis. Unfortunately, existing approaches to multilevel moderation have a variety of shortcomings, including conflated effects across levels of analysis and bias due to using observed cluster averages instead of latent variables (i.e., "random intercepts") to represent higher-level constructs. To overcome these problems and elucidate the nature of multilevel moderation effects, we introduce a multilevel structural equation modeling (MSEM) logic that clarifies the nature of the problems with existing practices and remedies them with latent variable interactions. This remedy uses random coefficients and/or latent moderated structural equations (LMS) for unbiased tests of multilevel moderation. We describe our approach and provide an example using the publicly available High School and Beyond data with Mplus syntax in Appendix. Our MSEM method eliminates problems of conflated multilevel effects and reduces bias in parameter estimates while offering a coherent framework for conceptualizing and testing multilevel moderation effects.

Keywords: multilevel modeling, moderation, interactions, latent variables, random slopes

Supplemental materials: http://dx.doi.org/10.1037/met0000052.supp

Researchers often theorize and test hypotheses of *moderation*, which occurs when the effect of an independent variable (a *focal predictor*) depends on the level of another variable (a *moderator*)—an *interaction*[1] effect (Cohen, Cohen, West, & Aiken, 2003). Researchers often also have *multilevel* or *hierarchically clustered* data, such as children nested within classrooms or employees in organizations, which are usually analyzed with multilevel modeling (MLM; e.g., Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). The result is that multilevel moderation tests are increasingly common. To facilitate these tests, multiple procedures have been proposed (e.g., Aguinis, Gottfredson, & Culpepper, 2013; Hofmann & Gavin, 1998; Preacher, Curran, & Bauer, 2006; Raudenbush, 1989a, 1989b; Raudenbush & Bryk, 1986). Unfortunately, there are various conceptual and statistical problems associated with these procedures.

These problems occur because most approaches to testing multilevel moderation do not separate lower- and higher-level effects into their orthogonal components, and instead conflate these effects by combining them into single coefficients (see Preacher, Zyphur, & Zhang, 2010). Across multiple fields, such conflation is known to cause model misspecification (see Hausman, 1978), resulting in conceptual and statistical problems. Conceptually, researchers (a) create theories and hypotheses that are insensitive to the different yet theoretically meaningful ways that moderation can occur, and (b) specify models reflecting this insensitivity. Statistically, researchers test moderation by (a) unknowingly constraining effects to equality across levels, and (b) introducing bias into estimates of moderation effects by not treating outcomes *and* predictors as latent variables at the levels of analysis stipulated in theory. As a result, researchers' theories are often tested with conflated and potentially biased parameter estimates, while theoretically meaningful moderation effects go undetected.

To address these issues our article has two goals. First, we advocate examining level-specific moderation effects, which are rarely considered but have important consequences for theory and practice. Second, to address the problem of multilevel interactions we recommend multilevel structural equation modeling (MSEM) for conceptualizing and estimating multilevel moderation. In so doing, we contribute to multilevel methods in two ways. First, our study is the first to use a latent variable approach to examine all

This article was published Online First December 14, 2015.

Kristopher J. Preacher, Department of Psychology and Human Development, Vanderbilt University; Zhen Zhang, Department of Management, Arizona State University; Michael J. Zyphur, Department of Management and Marketing, University of Melbourne.

An online appendix at http://quantpsy.org contains code for all of the models mentioned or illustrated in this article.

Correspondence concerning this article should be addressed to Kristopher J. Preacher, Department of Psychology and Human Development, Vanderbilt University, PMB 552, 230 Appleton Place, Nashville, TN 37203-5721. E-mail: kris.preacher@vanderbilt.edu

[1] We use the terms "interaction" and "moderation" interchangeably throughout.

possible multilevel moderation effects—effects that have different meanings and offer unique theoretical insights. Second, we illustrate two ways to test multilevel moderation in MSEM: traditional random coefficient prediction (RCP) for cross-level interactions; and latent moderated structural equations (LMS) for same-level interactions. We show how these methods can reduce bias and confusion in interpreting results. By extending LMS to the multilevel case, we also provide an example for researchers to correctly specify their models with RCP and/or LMS.

In what follows, we begin by outlining the logic of multilevel effect decomposition. We then treat the different types of multilevel moderation that can be hypothesized and tested. We map these hypotheses onto an MSEM framework that, compared with existing approaches, is clearer about multilevel effects, less prone to bias, and more versatile by allowing moderation hypotheses to be embedded in larger models. We then give an example of model specification and result interpretation, and conclude with extensions and caveats for our approach.

Decomposing Effects in MLM

In multilevel research, Level-1 (or L1) variables are measured at the lowest level of analysis (e.g., students); Level-2 (or L2) variables are measured at a second, higher level of analysis (e.g., classrooms). L1 variables can be partitioned into two parts, a part that varies only between L2 units (termed *between-cluster* or "B") and a part that varies only within L2 units (termed *within-cluster* or "W"; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). These components are population analogs of, respectively, cluster means (i.e., $y_{.j}$) and scores centered around the cluster means (i.e., $y_i = y_{ij} - y_{.j}$)—we use the "$.j$" subscript to denote variables that represent clusters' latent standings along an L1 variable, not an L2 variable.

Because L1 variables have B and W parts, they can also have B and W *effects* (Asparouhov & Muthén, 2006; Enders & Tofighi, 2007; Hedeker & Gibbons, 2006; Kreft, de Leeuw, & Aiken, 1995; Lüdtke, Marsh, Robitzsch, & Trautwein, 2011; Lüdtke et al., 2008; Mancl, Leroux, & DeRouen, 2000; Neuhaus & Kalbfleisch, 1998; Neuhaus & McCulloch, 2006). L2 variables have no W part and so are regarded as "only B" variables. Studies may also have "only W" variables, but we do not treat this case because it is rare and implies independent observations.

Between- and Within-Cluster Main Effects

Decomposing B and W effects is discussed by many authors (e.g., Cronbach, 1976; Curran & Bauer, 2011; Lüdtke et al., 2011, 2008; Raudenbush & Bryk, 2002). Their concerns are almost always with decomposing B and W parts of *main effects only* (i.e., separating effects of L1 predictors into B and W parts). Consider Equation 1, containing a single "conflated" effect of x_{ij} (i.e., a B and W effect as a single coefficient β_{1j}):

$$\begin{aligned} y_{ij} &= \beta_{0j} + \beta_{1j}x_{ij} + \varepsilon_{ij} \\ \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \\ y_{ij} &= \gamma_{00} + \gamma_{10}x_{ij} + u_{0j} + u_{1j}x_{ij} + \varepsilon_{ij} \end{aligned} \tag{1}$$

Here, the L1 variable x_{ij} permits only one effect on y_{ij} with mean γ_{10} and a B variance $\text{var}(u_{1j}) = \tau_{11}$. Yet, x_{ij} can be seen as the sum of W and B parts (x_i and $x_{.j}$). This alternative conceptualization replaces x_{ij} in Equation 1 with $(x_i + x_{.j})$, and decomposition yields the following separation of the slope of x_{ij} $(\gamma_{10} + u_{1j})$ into the two effects $\gamma^*_{01} + u^*_{1j}$ and $\gamma^*_{10} + u^*_{1j}$:

$$\begin{aligned} y_{ij} &= \gamma_{00} + \gamma_{10}x_{ij} + u_{0j} + u_{1j}x_{ij} + \varepsilon_{ij} \\ y_{ij} &= \gamma_{00} + \gamma^*_{10}x_i + \gamma^*_{01}x_{.j} + u_{0j} + u^*_{1j}x_i + u^*_{1j}x_{.j} + \varepsilon_{ij} \end{aligned} \tag{2}$$

The term γ^*_{10} is the effect of x_i (individuals' standings on x_{ij} relative to cluster j's mean) on y_{ij}, whereas γ^*_{01} is the *latent cluster mean's* effect of x_{ij} (not the observed cluster mean). Critically, γ^*_{01} and γ^*_{10} may have distinct substantive meanings—and hence very different magnitudes and signs—that are conflated as γ_{10}. Further, the random slope residual u^*_{1j} is multiplied not only by x_i, but also $x_{.j}$, which may be difficult to justify.[2]

The effects of level-specific components of variables may not always be of interest. At times, the overall effect of a predictor may be of interest. However, estimation of conflated effects combines W and B effects into a single estimate, and thus risks mischaracterizing both. Due to the possible differences in W and B effects, some authors advocate first decomposing these effects (e.g., Cronbach, 1976; Preacher et al., 2010), and we take this approach here.

It is important to recognize that $x_{.j}$ in Equation 2 is not a cluster mean; it is the latent standing of cluster j along x_{ij}. This distinction of means and latent standings of clusters is found in MLM literature stressing that, because clusters' standings on L1 variables are latent, cluster means measure clusters' standings with error (Lüdtke et al., 2011, 2008; Marsh et al., 2009). This causes bias by pulling $\hat{\gamma}^*_{01}$ toward $\hat{\gamma}^*_{10}$ (Preacher, Zhang, & Zyphur, 2011). Although large cluster sizes, for example, reduce the magnitude of this problem, to fully account for this bias requires MSEM, wherein latent B parts are used in lieu of means (Lüdtke et al., 2011, 2008; Preacher et al., 2010). In turn, testing moderation with latent components requires latent variable interactions to form product terms at two levels. Below, we treat these latent interactions, but first introduce the conceptual logic of moderation analyses at multiple levels.

Between- and Within-Cluster Interaction Effects

Some MLM literature has treated the decomposition of interactions (Enders & Tofighi, 2007; Hofmann & Gavin, 1998; Raudenbush, 1989a, 1989b; Raudenbush & Bryk, 1986; Ryu, 2015). This decomposition for moderation is more complex than for mediation because the product of two L1 predictors x_{ij} and z_{ij} $(x_{ij}z_{ij})$ should not be separated into B and W parts $(xz)_i$ and $(xz)_{.j}$ such that $(xz)_i + (xz)_{.j} = x_{ij}z_{ij}$. If this strategy were followed for L1 predictors x_{ij} and z_{ij} that were correlated at any value other than 0, the B and W parts of the product term $x_{ij}z_{ij}$ would be expanded as, respectively,

[2] The u^*_{1j} will differ from u_{1j}. The u^*_{1j} will be smaller on average than corresponding u_{1j}, and interpretable as a mix of two quantities: the departure of a given cluster's x_i slope from γ^*_{10}, and a cluster-specific weight permitting heteroscedastic variability in predicted values of y_{ij} that depends on the cluster mean $x_{.j}$.

$$(xz)_{.j} = E_j(x_{ij}z_{ij}) = x_{.j}z_{.j} + \mathrm{cov}_j(x_{.j}, z_{.j}) \tag{3}$$

and

$$(xz)_i = x_{ij}z_{ij} - E_j(x_{ij}z_{ij}) = x_{ij}z_{ij} - x_{.j}z_{.j} - \mathrm{cov}_j(x_{.j}, z_{.j}). \tag{4}$$

Thus, B and W standings along product terms are sensitive to the covariance of predictors that form the terms. Although $(xz)_i$ and $(xz)_{.j}$ sum to $x_{ij}z_{ij}$, using them as predictors does not lead to interpretable effects because researchers are not interested in the effects of product terms. Interest lies in the conditional slopes implied by the interaction terms, to which we now turn.

Multilevel Moderation Hypotheses

Methods for testing moderation in MLM borrow from single-level regression approaches (e.g., Aiken & West, 1991; Jaccard & Turrisi, 2003), translated into a multilevel context (Bauer & Curran, 2005; Preacher et al., 2006). For example, MLM literature shows how to include L1 × L1, L1 × L2, and L2 × L2 product terms in a multilevel model and plot and probe the effects. In some cases, product terms must be constructed prior to model specification and estimation by computing products of focal predictors and moderators, and including them in the data set (for L1 × L1 or L2 × L2 interactions). In other cases, a cross-level product term is implied if a random slope for an L1 predictor is regressed onto an L2 predictor—for L1 × L2 interactions, as we show below, this is a classic "slopes-as-outcomes" RCP model (Raudenbush & Bryk, 2002).

However, there is little discussion of how product terms in single- and multilevel contexts are incommensurable. In MLM, L1 predictors have B and W parts, so decomposing moderation effects at B and W levels is required but has not yet been treated comprehensively. To correct this oversight we now describe several possible types of multilevel moderation.

Moderation in 1 × (1 → 1) Designs

We begin by considering the following example (inspired by Kidwell, Mossholder, & Bennett, 1997), in which employees ($i = 1 \ldots n_j$) are nested within teams ($j = 1 \ldots J$).

> *Example A:* Researchers test whether the effect of employee job satisfaction (x_{ij}, the focal predictor) on employee courtesy (y_{ij}) varies across levels of employee organizational commitment (z_{ij}, the moderator). All three variables are assessed at the employee level (L1) rather than the team level (L2).

We denote this design 1 × (1 → 1), where the first "1" is the level at which the moderator is measured, the second "1" is the level at which the focal predictor is measured, and the last "1" is the level at which the outcome is measured. Using one currently popular strategy (Bauer & Curran, 2005; Preacher et al., 2006), the researcher would fit the following multilevel model:

$$\begin{aligned}
&\text{L1 equation:} && y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \beta_{2j}z_{ij} + \beta_{3j}x_{ij}z_{ij} + \varepsilon_{ij}\\
&\text{L2 equations:} && \begin{cases}\beta_{0j} = \gamma_{00} + u_{0j}\\ \beta_{1j} = \gamma_{10} + u_{1j}\\ \beta_{2j} = \gamma_{20} + u_{2j}\\ \beta_{3j} = \gamma_{30} + u_{3j}\end{cases}\\
&\text{Reduced-form:} && y_{ij} = \underbrace{\gamma_{00} + \gamma_{10}x_{ij} + \gamma_{20}z_{ij} + \gamma_{30}x_{ij}z_{ij}}_{\text{fixed part}}\\
&&& \quad + \underbrace{u_{0j} + u_{1j}x_{ij} + u_{2j}z_{ij} + u_{3j}x_{ij}z_{ij} + \varepsilon_{ij}}_{\text{random part}}
\end{aligned} \tag{5}$$

wherein the β_j terms are potentially random coefficients, the γ terms are fixed, ε_{ij} is a L1 residual, and u terms are L2 residuals. Including the product of interacting predictors ($x_{ij}z_{ij}$) in the model as a predictor enables the researcher to investigate moderation. The estimated slope $\hat{\gamma}_{30}$ is the interaction effect. If $\hat{\gamma}_{30}$ is statistically significant, the interaction is "probed and plotted" by assessing the significance of simple slopes of y_{ij} regressed on x_{ij} (the focal predictor) at conditional values of z_{ij} (the moderator) and by plotting conditional regression lines at these values (Aiken & West, 1991; Bauer & Curran, 2005; Preacher et al., 2006).

This practice is problematic because it does not clarify the many potential moderation hypotheses for x_{ij} and z_{ij} that may be substantively meaningful and can provide theoretical insights. To see the potential forms of moderation, consider Equation 6, which reformulates the model in Equation 5 by partitioning each L1 variable into B and W parts with separate slopes (letting $x_{ij} = x_i + x_{.j}$ and $z_{ij} = z_i + z_{.j}$, but recalling that the B parts will be treated as latent, as in Lüdtke et al., 2011, 2008; Marsh et al., 2009; and Preacher et al., 2010 as we illustrate later).

$$\begin{aligned}
&\text{L1 equation:} && y_{ij} = \beta_{0j} + \beta_{1j}x_i + \beta_{2j}z_i + \beta_{3j}x_iz_i + \varepsilon_{ij}\\
&\text{L2 equations:} && \begin{cases}\beta_{0j} = \gamma_{00} + \gamma_{01}x_{.j} + \gamma_{02}z_{.j} + \gamma_{03}x_{.j}z_{.j} + u_{0j}\\ \beta_{1j} = \gamma_{10} + \gamma_{11}z_{.j} + u_{1j}\\ \beta_{2j} = \gamma_{20} + \gamma_{21}x_{.j} + u_{2j}\\ \beta_{3j} = \gamma_{30} + u_{3j}\end{cases}\\
&\text{Reduced-form:} && y_{ij} = \underbrace{\gamma_{00} + \gamma_{01}x_{.j} + \gamma_{02}z_{.j} + \gamma_{10}x_i + \gamma_{20}z_i + \gamma_{03}x_{.j}z_{.j} + \gamma_{11}x_iz_{.j} + \gamma_{21}x_{.j}z_i + \gamma_{30}x_iz_i}_{\text{fixed part}}\\
&&& \quad + \underbrace{u_{0j} + u_{1j}x_i + u_{2j}z_i + u_{3j}x_iz_i + \varepsilon_{ij}}_{\text{random part}}
\end{aligned} \tag{6}$$

Note that parameters and residuals with the same symbols in Equations 5 and 6 have different interpretations. Unlike Equation 5, which has one interaction effect with fixed and random parts, the population model in Equation 6 has *four* interaction effects: a W interaction with fixed and random parts ($\gamma_{30} + u_{3j}$), two fixed cross-level interactions (γ_{11} and γ_{21}), and a B interaction (γ_{03}).[3] Thus, traditional analyses mask potentially important types of interactions.

[3] Any of these fixed effects could be treated as random at Level 3 if a third level exists.

Before continuing we make three points. First, not all possible interactions will always be of interest, and limiting the number of effects investigated will ease model estimation. Second, Equation 6 shows that the model in Equation 5 implies terms $u_{1j}x_{.j}$, $u_{2j}z_{.j}$, $u_{3j}x_iz_{.j}$, $u_{3j}x_{.j}z_i$, and $u_{3j}x_{.j}z_{.j}$. In brief, these terms allow the B effects of both predictors to vary across clusters. Although these terms are not explicit in models like Equation 5, similar terms are discussed in econometrics as capturing heteroscedasticity (e.g., Hildreth & Houck, 1968; Johnston, 1984; Wooldridge, 2002). Whatever their interpretation, these effects are not identifiable without model constraints and we do not discuss them further because we presume that they are rarely of interest. Third, we could augment Equation 6 to regress β_{1j} on $x_{.j}$ and regress β_{2j} on $z_{.j}$, essentially permitting cross-level interactions of the W and B parts of each variable with itself (i.e., the same variables). Such interactions may not be uncommon, but they are not our main interest.

Given the difference between Equations 5 and 6, three sensible moderation hypotheses are possible for scenarios like Example A, wherein x_{ij} is a focal predictor and z_{ij} is a moderator. These and subsequent prototypical hypotheses are listed in Table 1 for easy reference.

> *Hypothesis A1:* The W part of z_{ij} moderates the W effect of x_{ij} on y_{ij}, so that the effect of an employee's relative job satisfaction within a team on an employee's relative courtesy is moderated by the employee's relative organizational commitment in a team.

> *Hypothesis A2:* The B part of z_{ij} moderates the W effect of x_{ij} on y_{ij},[4] meaning the effect of an employee's relative job satisfaction within a team on an employee's relative courtesy is moderated by a team's collective organizational commitment.

> *Hypothesis A3:* The B part of z_{ij} moderates the B effect of x_{ij} on y_{ij}, meaning that the effect of a team's collective job satisfaction on collective employee courtesy is moderated by the team's collective organizational commitment.

Examining only $\hat{\gamma}_{30}$ in Equation 5 conflates coefficients that should be separately estimated, as in Equation 6—it may be that the interaction effect in Hypothesis A1 is positive and significant, but the effect in Hypothesis A2 is negative and nonsignificant, and so on. Also, the single conflated coefficient $\hat{\gamma}_{30}$ in Equation 5 can be influenced by intraclass correlations (ICC) and sample size, but this is not true for the coefficients in Hypotheses A1–A3 (see Cronbach & Snow, 1977).

Table 1
Possible Two-Way Interaction Hypotheses in a Two-Level Model

Hypothesis	Notation	Design
A1	$z_i \times (x_i \rightarrow y_{ij})$	$1 \times (1 \rightarrow 1)$
A2	$z_{.j} \times (x_i \rightarrow y_{ij})$	$1 \times (1 \rightarrow 1)$
A3	$z_{.j} \times (x_{.j} \rightarrow y_{ij})$	$1 \times (1 \rightarrow 1)$
B1	$z_j \times (x_i \rightarrow y_{ij})$	$2 \times (1 \rightarrow 1)$
B2	$z_j \times (x_{.j} \rightarrow y_{ij})$	$2 \times (1 \rightarrow 1)$
C	$z_j \times (x_j \rightarrow y_{ij})$	$2 \times (2 \rightarrow 1)$
D	$z_{.j} \times (x_j \rightarrow y_{ij})$	$1 \times (2 \rightarrow 1)$

Note. The numbers in the Design column show the level at which each variable was measured (rather than conceptualized and analyzed).

Moderation in 2 × (1 → 1) Designs

> *Example B:* Researchers test whether the effect of job satisfaction (x_{ij}, the focal predictor) on courtesy (y_{ij}) is moderated by team cohesiveness (z_j, the moderator), measured at L2.

Because the moderator is at L2 and both the focal predictor and outcome are at L1, we denote this a 2 × (1 → 1) design. This yields the following reduced-form multilevel model:

$$y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + \gamma_{01}z_j + \gamma_{11}x_{ij}z_j + u_{0j} + u_{1j}x_{ij} + \varepsilon_{ij} \tag{7}$$

Most MLM literature tests $\hat{\gamma}_{11}$ to assess moderation, but γ_{10} and γ_{11} are conflated coefficients due to x_{ij} varying at W and B levels. If x_{ij} is partitioned into level-specific parts we obtain:

$$\begin{aligned}
&\text{L1 equation:} && y_{ij} = \beta_{0j} + \beta_{1j}x_i + \varepsilon_{ij}\\
&\text{L2 equations:} && \begin{cases}\beta_{0j} = \gamma_{00} + \gamma_{01}x_{.j} + \gamma_{02}z_j + \gamma_{03}x_{.j}z_j + u_{0j}\\ \beta_{1j} = \gamma_{10} + \gamma_{11}z_j + u_{1j}\end{cases}\\
&\text{Reduced-form:} && y_{ij} = \underbrace{\gamma_{00} + \gamma_{01}x_{.j} + \gamma_{02}z_j + \gamma_{03}x_{.j}z_j + \gamma_{10}x_i + \gamma_{11}x_iz_j}_{\text{fixed part}}\\
& && \quad + \underbrace{u_{0j} + u_{1j}x_i + \varepsilon_{ij}}_{\text{random part}}
\end{aligned} \tag{8}$$

(As before, terms with the same symbols in Equations 7 and 8 have different meanings.) Equation 8 shows that there are *two* interaction hypotheses in 2 × (1 → 1) designs, which may be stated as:

> *Hypothesis B1:* The B variable z_j moderates the W effect of x_{ij} on y_{ij}, such that the effect of an employee's relative job satisfaction in a team on employee courtesy is moderated by the team's cohesiveness.

> *Hypothesis B2:* The B variable z_j moderates the B effect of x_{ij} on y_{ij}, such that the effect of a team's collective job satisfaction on employee courtesy is moderated by the team's cohesiveness.

The interaction effect in Equation 7 is often termed *cross-level*. Yet, only Hypothesis B1's slope (γ_{11} in Equation 8) is truly cross-level in that an L2 moderator influences a W slope (Aguinis et al., 2013; Mathieu, Aguinis, Culpepper, & Chen, 2012). Thus, common practice conflates the two γs that a researcher presumably would prefer to examine separately (for Hypotheses B1 and B2), and different conclusions for sign, magnitude, and practical significance could be drawn for Hypotheses B1 and B2 had each been tested separately. Further, although decomposing L1 predictors in

[4] We omit the potential hypothesis that the B part of x_{ij} could moderate the W effect of z_{ij} because this hypothesis is identical in form to Hypothesis A2 (after switching variable labels), and would require identical methods. Both effects may exist in the same model; for examples, see van Yperen and Snijders (2000).

cross-level interactions has been suggested (Aguinis et al., 2013; Cronbach & Webb, 1975; Enders & Tofighi, 2007; Hofmann & Gavin, 1998; Raudenbush, 1989a, 1989b; Raudenbush & Bryk, 1986), the popular method of centering L1 predictors at the cluster mean and including this mean as a predictor biases estimates of B effects (Preacher et al., 2011).

Moderation in 2 × (2 → 1) Designs

Example C: Researchers test whether the effect of team salary range (x_j, the focal predictor) on employee courtesy (y_{ij}) is moderated by team cohesiveness (z_j, the moderator). Both x_j and z_j are L2 variables and team cohesiveness is assessed by external raters.

Because the focal predictor and the moderator are assessed at L2, we denote this a 2 × (2 → 1) design, yielding the following reduced-form multilevel model:

$$y_{ij} = \gamma_{00} + \gamma_{01}x_j + \gamma_{02}z_j + \gamma_{03}x_jz_j + u_{0j} + \varepsilon_{ij} \tag{9}$$

MLM literature appropriately tests $\hat{\gamma}_{03}$ to examine moderation—x_j and z_j do not vary within clusters, so they cannot be partitioned into B and W parts. The only moderation hypothesis is:

Hypothesis C: The B variable z_j moderates the B effect of x_j on y_{ij}, meaning that the effect of a team's salary range on employee courtesy is moderated by the team's cohesiveness.

Yet, if team cohesiveness is rated by team members, using team means to calculate x_jz_j biases estimates of $\hat{\gamma}_{03}$. In such cases, cluster means should be treated as latent (as in Hypothesis A3).

Moderation in 1 × (2 → 1) Designs

Example D: Researchers test whether the effect of team diversity (x_j, the focal predictor) on employee creativity (y_{ij}) is moderated by the employee's openness to experience (z_{ij}, the moderator). x_j is a L2 variable and y_{ij} and z_{ij} are assessed at L1.

Because the focal predictor is L2 and the moderator of this B effect is at L1, we denote this design a 1 × (2 → 1) design. This yields the following reduced-form multilevel model:

$$y_{ij} = \gamma_{00} + \gamma_{10}z_{ij} + \gamma_{01}x_j + \gamma_{11}z_{ij}x_j + u_{0j} + u_{1j}z_{ij} + \varepsilon_{ij} \tag{10}$$

In prior work using this design, researchers often tested a cross-level interaction using $\hat{\gamma}_{11}$ from Equation 7. Yet, this conflates the B and W parts of z_{ij} if the proper centering method is not used. More importantly, it overlooks how the hypothesis concerns not whether x_j moderates the slope of y_{ij} regressed on z_{ij}, but instead whether the B part of z_{ij} moderates the B effect of x_j on y_{ij}. Partitioning Equation 10 makes this point:

$$\begin{aligned}
&\text{L1 equation:} && y_{ij} = \beta_{0j} + \beta_{1j}z_i + \varepsilon_{ij}\\
&\text{L2 equations:} && \begin{cases}\beta_{0j} = \gamma_{00} + \gamma_{01}z_{.j} + \gamma_{02}x_j + \gamma_{03}z_{.j}x_j + u_{0j}\\ \beta_{1j} = \gamma_{10} + \gamma_{11}x_j + u_{1j}\end{cases}\\
&\text{Reduced-form:} && y_{ij} = \underbrace{\gamma_{00} + \gamma_{01}z_{.j} + \gamma_{02}x_j + \gamma_{03}z_{.j}x_j + \gamma_{10}z_i + \gamma_{11}z_ix_j}_{\text{fixed part}}\\
& && \quad + \underbrace{u_{0j} + u_{1j}z_i + \varepsilon_{ij}}_{\text{random part}}
\end{aligned} \tag{11}$$

As in Example B, there are two interactions in Equation 11, but probably only one is of interest: γ_{03}, the moderation of x_j's B effect by the B part of z_{ij}. The corresponding hypothesis is:

Hypothesis D: The B part of z_{ij} moderates the B effect of x_j on y_{ij}, so that the effect of team diversity on employee creativity is moderated by the team's collective openness to experience.

The other interaction in Equation 11, γ_{11}, is cross-level with the predictor (x_j) moderating the W effect of the moderator (z_{ij}). Thus, the roles of moderator and predictor switch if we interpret γ_{11}, rendering the research question substantively different from Example D. Example D does not concern a cross-level interaction, but rather a B interaction of L1 and L2 predictors.

Summary

To summarize, current practices for testing multilevel moderation present multiple problems. First, they do not help researchers hypothesize and test all moderation relationships that may exist (e.g., Hypotheses A1, A2, A3, B1, B2, and D). Second, many published interaction effects are "uninterpretable blends" of multiple and substantively unique coefficients that may differ in direction, magnitude, and interpretation across levels (Cronbach, 1976, p. 9.20). Further, suggested solutions for this second problem treat means of L1 variables as observed, yielding biased effect estimates, with bias increasing as a predictor's ICC and cluster size decrease in the presence of different effects across levels (Lüdtke et al., 2011, 2008).

Below, we describe how to solve these problems by conceptualizing and testing all possible interactions involving the W and B parts of L1 variables with each other or with L2 variables, and/or among L2 variables (as in Examples A–D above). This requires an approach that (a) decomposes observed L1 variables into their latent B and W parts, and (b) allows multiplying these latent variables or components of variables to form product terms. To do this, we connect literature on observed variable interactions in MLM, the decomposition of B and W effects in MLM, and latent variable interactions, all within an MSEM framework.

An MSEM Framework for Assessing Multilevel Moderation

Current MSEM methods build upon initial work by Schmidt (1969). Major developments in specification and estimation followed (see Goldstein & McDonald, 1988; McDonald, 1993, 1994; McDonald & Goldstein, 1989; Muthén, 1989, 1990, 1991, 1994;

Muthén & Satorra, 1995). Early MSEM approaches decomposed observed data into B and pooled W covariance matrices, fitting separate B and W models with the multigroup function of SEM software, which did not allow random coefficients, missing data, or unbalanced cluster sizes. Recent MSEM methods overcome these limitations (e.g., Chou, Bentler, & Pentz, 2000; Jedidi & Ansari, 2001; Muthén & Asparouhov, 2009; Raudenbush & Sampson, 1999; Skrondal & Rabe-Hesketh, 2004). Each method has benefits and drawbacks, but Muthén and Asparouhov's (2009) has algorithms that make complex models tractable in Mplus (Muthén & Muthén, 1998–2014). Benefits of this MSEM method are: it decomposes variables and effects into B and W parts; it allows outcome variables at L2 (unlike MLM); it treats the B part of L1 variables as latent (Lüdtke et al., 2011, 2008); and it allows latent variable interactions for testing multilevel moderation.

All of this is done by adding random coefficients to the familiar SEM equations, treating a subset of parameters as randomly (co)varying across clusters and modeling the (co)variances with a B structural model. The equations for MSEM can be shown as follows:

$$\mathbf{y}_{ij} = \mathbf{\Lambda}_j \boldsymbol{\eta}_{ij} \tag{12}$$

$$\boldsymbol{\eta}_{ij} = \boldsymbol{\alpha}_j + \mathbf{B}_j \boldsymbol{\eta}_{ij} + \boldsymbol{\zeta}_{ij} \tag{13}$$

$$\boldsymbol{\eta}_j = \boldsymbol{\mu} + \boldsymbol{\beta} \boldsymbol{\eta}_j + \boldsymbol{\zeta}_j \tag{14}$$

Equations 12 and 13 are similar to measurement and structural equations (respectively) commonly seen in SEM. A j subscript on matrices in Equations 12 and 13 indicates that some elements may vary across clusters. In these equations, $\mathbf{\Lambda}_j$ is a matrix of 0 and 1 loadings[5] linking observed variables in $\mathbf{y}_{ij}$ to latent variables in $\boldsymbol{\eta}_{ij}$, $\boldsymbol{\eta}_{ij}$ is a vector of all latent variables from each level, $\boldsymbol{\alpha}_j$ is a vector of random coefficients for the latent variables, $\mathbf{B}_j$ is a matrix of structural effects relating latent variables, $\boldsymbol{\zeta}_{ij}$ contains W structural residuals, $\boldsymbol{\eta}_j$ is a vector of all random coefficients (including random slopes), $\boldsymbol{\mu}$ contains intercepts or means of all of the random coefficients, $\boldsymbol{\beta}$ contains effects relating random coefficients to each other, and $\boldsymbol{\zeta}_j$ contains B structural residuals—it is customary to permit the residuals of random effects not otherwise related to freely covary. Of course, as with all SEM-based analyses, strong theory rather than merely model-based estimation is required to enhance the veracity of causal inferences.

We now describe tests of multilevel moderation using two methods that treat the B parts of L1 variables as latent: *random coefficient prediction* (RCP) and *latent moderated structural equations* (LMS). Whereas RCP is already well known as "slopes-as-outcomes" MLM, LMS is a new method for multilevel moderation. LMS is an important tool because RCP is suited for cross-level interactions that include latent B parts of L1 variables, but LMS is preferable for same-level interactions for multiple reasons. For example, because the random slope of a L1 predictor is an L2 variable, it cannot be expressed as a function of another L1 predictor (to test Hypothesis A1), so RCP is applicable only to interactions with at least one L2 variable or B component of an L1 variable. Alternatively, LMS is applicable in all situations we describe.

Random Coefficient Prediction (RCP) Method

In RCP a random slope is predicted by a moderator. The slope's residual variance can be fixed to zero (or a small value to assist in estimation), so it need not be "random" while still varying as a function of predictors, as in typical cross-level interactions with L1 predictors and L2 moderators. Consider Equation 7 from Example B, expressed with separate L1 and L2 equations:

$$\begin{aligned}
&\text{L1 equation:} && y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \varepsilon_{ij}\\
&\text{L2 equations:} && \begin{cases}\beta_{0j} = \gamma_{00} + \gamma_{01}z_j + u_{0j}\\ \beta_{1j} = \gamma_{10} + \gamma_{11}z_j + u_{1j}\end{cases}\\
&\text{Reduced-form:} && y_{ij} = \underbrace{\gamma_{00} + \gamma_{10}x_{ij} + \gamma_{01}z_j + \gamma_{11}x_{ij}z_j}_{\text{fixed part}} + \underbrace{u_{0j} + u_{1j}x_{ij} + \varepsilon_{ij}}_{\text{random part}}
\end{aligned} \tag{15}$$

In the L2 slope equation, the random slope β_{1j} is regressed on the L2 moderator z_j, implying a reduced-form equation with a product term $x_{ij}z_j$ multiplied by γ_{11}. In Equation 15, β_{1j} is conflated because x_{ij} has B and W variance (and therefore B and W effects). However, x_{ij} can be split into W and B parts, yielding Equation 8 in Example B. In the reduced-form of Equation 8, the W slope of x_i is moderated by the B moderator z_j, addressing Hypothesis B1. This is a truly cross-level interaction—one predictor has strictly W variance and the other strictly B variance.

To extend this logic to MSEM, consider how MSEM allows testing Hypothesis A2 in a $1 \times (1 \rightarrow 1)$ design, as in Example A. The measurement equation (see Equation 12) is:

$$\mathbf{y}_{ij} = \begin{bmatrix} x_{ij} \\ z_{ij} \\ y_{ij} \end{bmatrix} = \mathbf{\Lambda}\boldsymbol{\eta}_{ij} = \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} \eta_{xi} \\ \eta_{zi} \\ \eta_{yi} \\ -- \\ \eta_{xj} \\ \eta_{zj} \\ \eta_{yj} \end{bmatrix} \tag{16}$$

This equation links observed variables to latent components in $\boldsymbol{\eta}_{ij}$ that vary strictly at W or B levels. The vector $\mathbf{y}_{ij}$ contains all observed L1 variables for member i of cluster j. The loadings in $\mathbf{\Lambda}$ are fixed to 1 or 0 to associate each observed L1 variable with one or both of two level-specific components. For example, Equation 16 implies that the observed variable z_{ij} is expressible as the sum of the W latent component η_{zi} and the B latent component η_{zj}.

Equation 17 expresses the L1 elements of the latent vector $\boldsymbol{\eta}_{ij}$ as functions of fixed and random coefficients and the L1 parts of other L1 predictors. The L2 elements of $\boldsymbol{\eta}_{ij}$ are associated with L2

[5] These loadings are used merely to decompose variables into latent B and W parts. Although MSEM does permit latent variables with multiple indicators, we focus here on multilevel path analysis for simplicity.

parts which are modeled in Equation 18 (denoted α). To test Hypothesis A2, the W part of x_{ij} must have a random slope that is regressed on the B part of z_{ij} (the residual variance of this slope could be constrained to 0). In Equation 17, this random slope is $B_{y.xj}$ (similar to β_{1j} in Equation 6). The W effect of z_{ij} is not included in Equation 17, but could be added if desired.

$$\boldsymbol{\eta}_{ij}=\boldsymbol{\alpha}_j+\mathbf{B}_j\boldsymbol{\eta}_{ij}+\boldsymbol{\zeta}_{ij}$$
$$=\begin{bmatrix}\eta_{xi}\\ \eta_{zi}\\ \eta_{yi}\\ --\\ \eta_{xj}\\ \eta_{zj}\\ \eta_{yj}\end{bmatrix}=\begin{bmatrix}0\\0\\0\\ ---\\ \alpha_{\eta xj}\\ \alpha_{\eta zj}\\ \alpha_{\eta yj}\end{bmatrix}+\left[\begin{array}{ccc:ccc}0&0&0&0&0&0\\0&0&0&0&0&0\\B_{y.xj}&0&0&0&0&0\\ \hdashline 0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{array}\right]\begin{bmatrix}\eta_{xi}\\ \eta_{zi}\\ \eta_{yi}\\ --\\ \eta_{xj}\\ \eta_{zj}\\ \eta_{yj}\end{bmatrix}+\begin{bmatrix}\zeta_{xi}\\ \zeta_{zi}\\ \zeta_{yi}\\ --\\ 0\\0\\0\end{bmatrix} \tag{17}$$

In Equation 18 the four random coefficients from Equation 17 are stacked into a vector $\boldsymbol{\eta}_j$ and expressed as functions of means (μ) and of each other:

$$\boldsymbol{\eta}_j=\boldsymbol{\mu}+\boldsymbol{\beta}\boldsymbol{\eta}_j+\boldsymbol{\zeta}_j$$
$$=\begin{bmatrix}B_{y.xj}\\ \alpha_{\eta xj}\\ \alpha_{\eta zj}\\ \alpha_{\eta yj}\end{bmatrix}=\begin{bmatrix}\mu_{Byxj}\\ \mu_{\alpha\eta xj}\\ \mu_{\alpha\eta zj}\\ \mu_{\alpha\eta yj}\end{bmatrix}+\begin{bmatrix}0&0&\beta_{y.xz}&0\\0&0&0&0\\0&0&0&0\\0&\beta_{y.x}&\beta_{y.z}&0\end{bmatrix}\begin{bmatrix}B_{y.xj}\\ \alpha_{\eta xj}\\ \alpha_{\eta zj}\\ \alpha_{\eta yj}\end{bmatrix}+\begin{bmatrix}\zeta_{Byxj}\\ \zeta_{\alpha\eta xj}\\ \zeta_{\alpha\eta zj}\\ \zeta_{\alpha\eta yj}\end{bmatrix} \tag{18}$$

In Equation 18, $\beta_{y.x}$ and $\beta_{y.z}$ are conditional main effects of the B parts of x_{ij} and z_{ij} (similar to γ_{01} and γ_{02} in Equation 6), μ_{Byxj} is the mean of the random slope of the W part of x_{ij} (similar to γ_{10} in Equation 6), and $\beta_{y.xz}$ is the interaction effect of the W part of x_{ij} with the B part of z_{ij} (similar to γ_{11} in Equation 6)—this latter effect is the B part of z_{ij} moderating the W effect of x_{ij} on y_{ij}. Hypothesis B1 can be tested using a similar model that omits rows and columns for the W part of z_{ij}.

However, it is less well known that RCP can be used to test moderation in single-level models by regressing a dependent variable on a predictor with a random slope, and predicting the slope with a moderator (Muthén, 2011; Muthén & Muthén, 1998–2014, example 3.9; Yuan, Cheng, & Maxwell, 2014)—in single-level models, the slope's residual variance is constrained to zero. This approach can be used to model interactions corresponding to Hypotheses A3, B2, C, and D, but we do not explore it further because of bias, efficiency, and convergence problems that we describe later. Instead, we note that RCP is very useful for modeling cross-level interactions, but LMS can be used for cross- and same-level interactions, which we treat in the next section.

To illustrate our models we use a variant of SEM path diagrams (as in Muthén & Muthén, 1998–2014; Preacher, 2011; Preacher et al., 2011, 2010). As in path diagrams, circles are latent variables (or latent level-specific components), single-headed arrows are path coefficients and loadings, double-headed arrows are variances and covariances, triangles are unit constants for means and intercepts, and unattached arrows indicate residuals. Our variant of this convention places observed variables horizontally in the center of the diagram, the W model is below the center, and the B model is above the center. Random slopes have a darkened circle on the slope in question and appear as latent variables in the B model. In Figure 1 Panel A, a B1 hypothesis (cross-level interaction) is shown as the effect of the L2 moderator z_j on a random slope s_{1j}. The slope is treated as random so that it can be regressed onto z_j, but its residual variance and the covariance with y_j are fixed to zero for simplicity—these constraints may be relaxed in practice.

Latent Moderated Structural Equation (LMS) Method

Interactions are often tested by computing and then including products as predictors. Products are implicit in RCP, but they can be made explicit by computing a product involving at least one random B part of an L1 variable as a latent predictor. This cannot be done directly because it requires latent interactions. To solve this problem in the single-level case, multiple approaches have been suggested for Latent × Latent and Latent × Observed variable interactions.

Some approaches approximate latent × latent interactions with products of latent variable indicators (e.g., Jöreskog & Yang, 1996; Kenny & Judd, 1984). Leite and Zuo (2011) extended two such methods to MSEM: the *mean-centered unconstrained* (Marsh, Wen, & Hau, 2004) and *residual-centered unconstrained* (Little, Bovaird, & Widaman, 2006) approaches. When applied to multilevel data, both methods involve treating products of cluster means as indicators, but these are unreliable when ICCs, cluster sizes, and the correlation of the cluster-average variables are small. The approach we use in MSEM is superior in terms of bias and efficiency, and is simple in terms of specification: *latent moderated structural equations* (LMS;[6] Klein & Moosbrugger, 2000). Although meant to be a single-level approach, LMS can be applied to multilevel data by creating latent interactions among random coefficients. To explain, Klein and Moosbrugger (2000) and Schermelleh-Engel, Klein, and Moosbrugger (1998) proposed LMS for modeling two-way interactions between latent and/or measured variables. Consider a standard measurement model, in which measured "indicator" variables in the mean-centered vector $\mathbf{y}_i$ (containing indicators of latent predictors and latent outcomes) are related to latent variables $\boldsymbol{\eta}_i$ by loadings in $\boldsymbol{\Lambda}$, and $\boldsymbol{\delta}_i$ is a vector of residuals:

$$\mathbf{y}_i=\boldsymbol{\Lambda}\boldsymbol{\eta}_i+\boldsymbol{\delta}_i \tag{19}$$

[6] We use the abbreviation LMS to refer to models estimated using both the traditional LMS estimator, which employed ML estimation via the EM algorithm and a mixture of normal distributions to approximate the nonnormality of latent products, and the newer quasi-maximum likelihood (QML) estimator, which approximates the likelihood function by accommodating nonnormal product distributions with mixtures of normal distributions. Both methods are based on the same model specification, but only the mixture-based LMS method is included in Mplus. See Kelava, Moosbrugger, Dimitruk, and Schermelleh-Engel (2008), Kelava et al. (2011), Klein & Muthén (2007), and Moosbrugger, Schermelleh-Engel, Kelava, and Klein (2009) for further details, advantages, and disadvantages of these methods.

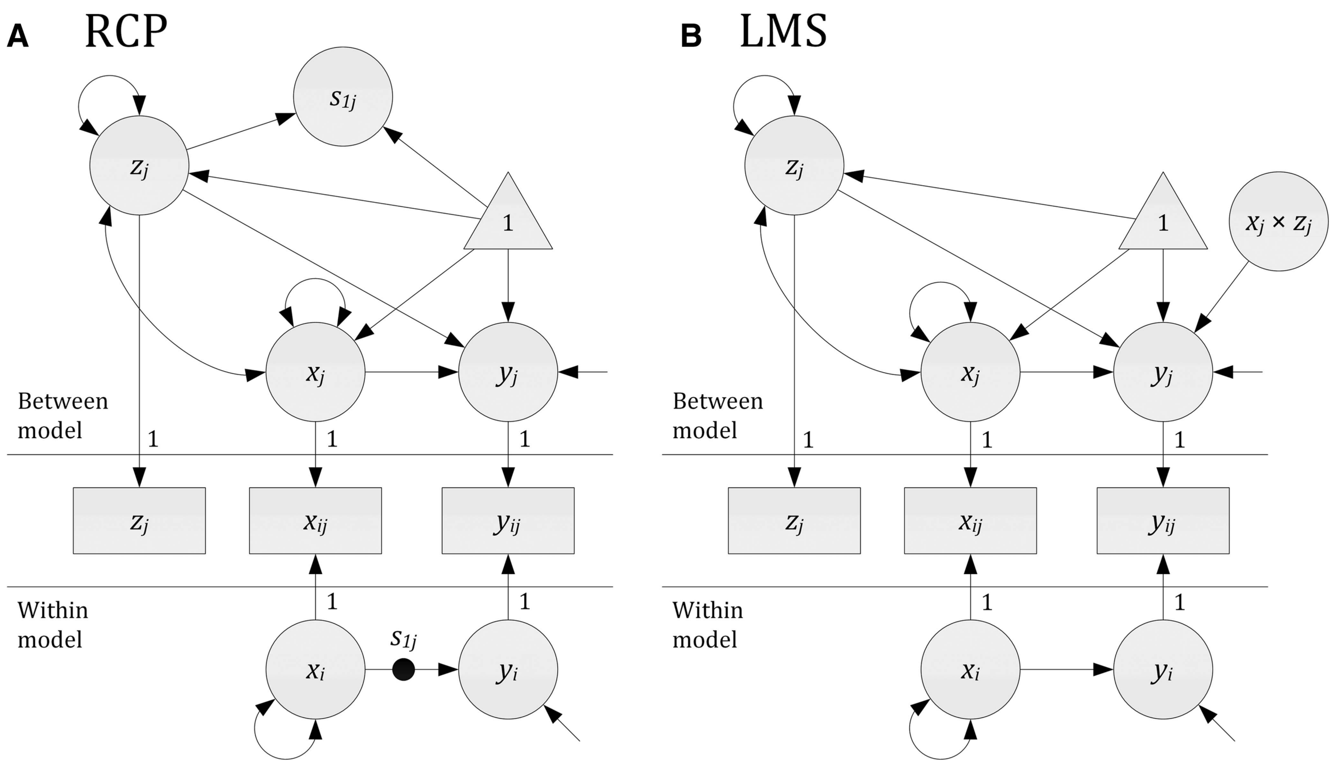

Figure 1. Panel A: Path diagram for a 2 × (1 → 1) design testing a cross-level interaction effect; z_j predicts s_{1j}, the random slope for the W component of x_{ij}. Panel B: Path diagram for a 2 × (1 → 1) design testing the interaction of z_j and the B component of x_{ij}. RCP = random coefficient prediction; LMS = latent moderated structural equations.

Unlike other strategies for assessing latent interactions, LMS does not rely on products of indicators, instead directly representing latent interactions in the structural equation:

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\eta}_i + \mathbf{j}\boldsymbol{\eta}_i'\boldsymbol{\Omega}\boldsymbol{\eta}_i + \boldsymbol{\zeta}_i \tag{20}$$

where $\boldsymbol{\alpha}$ contains intercepts, $\boldsymbol{\zeta}_i$ the structural regression residuals, $\mathbf{B}$ the conditional main effects of latent variables on other latent variables, and $\boldsymbol{\Omega}$ the quadratic or two-way interactions of latent variables. $\boldsymbol{\Omega}$ is upper-triangular with a zero diagonal. $\mathbf{j}$ is a vector that assigns these effects to one latent outcome in $\boldsymbol{\eta}_i$. Because observed indicator distributions depend on the nonnormal distribution of latent products, the joint distribution of the indicators is approximated with a continuous normal mixture. Interactions can be tested using either Wald or likelihood ratio tests (Klein & Moosbrugger, 2000), but the latter is more accurate in small samples and is scale-independent (Cham, West, Ma, & Aiken, 2012; Kelava et al., 2011). Mplus implements this approach and it performs well in simulations (see Klein & Muthén, 2007).

To see how LMS models interactions in MSEM, consider a model for Hypotheses A1–A3 in a 1 × (1 → 1) design. The measurement equation is as in Equation 16, linking observed variables to latent parts in $\boldsymbol{\eta}_{ij}$ that vary strictly at W or B levels. Equation 21 expands Equation 20 with product terms for LMS, expressing the W elements in $\boldsymbol{\eta}_{ij}$ as functions of the W parts of other L1 predictors as well as fixed and random coefficients. B elements of $\boldsymbol{\eta}_{ij}$ are linked to B components (denoted α), modeled in Equation 22. To test Hypotheses A1–A3, we modify Equation 20 to include a term reflecting regression of η_{yij} on the products of latent components of $\boldsymbol{\eta}_{ij}$. Specifically, $\boldsymbol{\eta}_{ij}'\boldsymbol{\Omega}_j\boldsymbol{\eta}_{ij}$ is a scalar sum, and $\mathbf{j}$ is a vector that assigns this sum to only one element of the latent vector $\boldsymbol{\eta}_{ij}$.[7]

$$\begin{aligned}
\boldsymbol{\eta}_{ij} &= \boldsymbol{\alpha}_j + \mathbf{B}_j\boldsymbol{\eta}_{ij} + \mathbf{j}\boldsymbol{\eta}_{ij}'\boldsymbol{\Omega}_j\boldsymbol{\eta}_{ij} + \boldsymbol{\zeta}_{ij} \\
&= \begin{bmatrix} \eta_{xi} \\ \eta_{zi} \\ \eta_{yi} \\ -- \\ \eta_{xj} \\ \eta_{zj} \\ \eta_{yj} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ --- \\ \alpha_{\eta xj} \\ \alpha_{\eta zj} \\ \alpha_{\eta yj} \end{bmatrix} + \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ B_{y.xj} & B_{y.zj} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} \eta_{xi} \\ \eta_{zi} \\ \eta_{yi} \\ \eta_{xj} \\ \eta_{zj} \\ \eta_{yj} \end{bmatrix} \\
&+ \begin{bmatrix} 0 \\ 0 \\ 1 \\ - \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \eta_{xi} \\ \eta_{zi} \\ \eta_{yi} \\ -- \\ \eta_{xj} \\ \eta_{zj} \\ \eta_{yj} \end{bmatrix}' \left[\begin{array}{ccc|ccc} 0 & B_{y.xizi} & 0 & 0 & B_{y.xizj} & 0 \\ 0 & 0 & 0 & B_{y.xjzi} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} \eta_{xi} \\ \eta_{zi} \\ \eta_{yi} \\ -- \\ \eta_{xj} \\ \eta_{zj} \\ \eta_{yj} \end{bmatrix} + \begin{bmatrix} \zeta_{xi} \\ \zeta_{zi} \\ \zeta_{yi} \\ -- \\ 0 \\ 0 \\ 0 \end{bmatrix}
\end{aligned} \tag{21}$$

[7] If multiple elements of $\boldsymbol{\eta}_{ij}$ are functions of products, multiple $\mathbf{j}\boldsymbol{\eta}_{ij}'\boldsymbol{\Omega}_j\boldsymbol{\eta}_{ij}$ terms can be entered in Equation 21. In Mplus, path coefficients associated with latent products are reported in the $\mathbf{B}_j$ and $\boldsymbol{\beta}$ matrices.

In Equation 21, $B_{y.xj}$ and $B_{y.zj}$ are slopes for the W parts of x_{ij} and z_{ij}. $B_{y.xizi}$ is the effect of the product of the W parts of x_{ij} and z_{ij}—the W interaction effect (Hypothesis A1). $B_{y.xizj}$ and $B_{y.xjzi}$ are cross-level interaction effects—respectively, the interaction of the W part of x_{ij} and B part of z_{ij}, and the interaction of the B part of x_{ij} and W part of z_{ij} (Hypothesis A2). Equation 22 stacks all random coefficients from Equation 21 in a vector $\boldsymbol{\eta}_j$, expressed as a function of means (μ):

$$
\begin{aligned}
\boldsymbol{\eta}_j &= \boldsymbol{\mu} + \boldsymbol{\beta\eta}_j + \mathbf{j}\boldsymbol{\eta}_j'\boldsymbol{\omega\eta}_j + \boldsymbol{\zeta}_j \\
&= \begin{bmatrix} B_{y.xj} \\ B_{y.zj} \\ B_{y.xizi} \\ B_{y.xizj} \\ B_{y.xjzi} \\ \alpha_{\eta xj} \\ \alpha_{\eta zj} \\ \alpha_{\eta yj} \end{bmatrix}
= \begin{bmatrix} \mu_{Byxj} \\ \mu_{Byzj} \\ \mu_{Byxizi} \\ \mu_{Byxizj} \\ \mu_{Byxjzi} \\ \mu_{\alpha\eta xj} \\ \mu_{\alpha\eta zj} \\ \mu_{\alpha\eta yj} \end{bmatrix}
+ \begin{bmatrix} 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&\beta_{y.x}&\beta_{y.z}&0 \end{bmatrix}
\begin{bmatrix} B_{y.xj} \\ B_{y.zj} \\ B_{y.xizi} \\ B_{y.xizj} \\ B_{y.xjzi} \\ \alpha_{\eta xj} \\ \alpha_{\eta zj} \\ \alpha_{\eta yj} \end{bmatrix} \\
&+ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}
\begin{bmatrix} B_{y.xj} \\ B_{y.zj} \\ B_{y.xizi} \\ B_{y.xizj} \\ B_{y.xjzi} \\ \alpha_{\eta xj} \\ \alpha_{\eta zj} \\ \alpha_{\eta yj} \end{bmatrix}'
\left[\begin{array}{ccccc|ccc} 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ \hline 0&0&0&0&0&0&\beta_{y.xz}&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{array}\right]
\begin{bmatrix} B_{y.xj} \\ B_{y.zj} \\ B_{y.xizi} \\ B_{y.xizj} \\ B_{y.xjzi} \\ \alpha_{\eta xj} \\ \alpha_{\eta zj} \\ \alpha_{\eta yj} \end{bmatrix}
+ \begin{bmatrix} \zeta_{Byxj} \\ \zeta_{Byzj} \\ \zeta_{Byxizi} \\ 0 \\ 0 \\ \zeta_{\alpha\eta xj} \\ \zeta_{\alpha\eta zj} \\ \zeta_{\alpha\eta yj} \end{bmatrix}
\end{aligned}
\tag{22}
$$

Equations 16, 21, and 22 imply the following equations for x_{ij}, z_{ij}, and y_{ij}, with B and W parts as:

$$
x_{ij} = \eta_{xj} + \eta_{xi} = \underbrace{\mu_{\alpha\eta xj} + \zeta_{\alpha\eta xj}}_{\text{B part of } x_{ij}} + \underbrace{\zeta_{xi}}_{\substack{\text{W part} \\ \text{of } x_{ij}}}
\tag{23}
$$

$$
z_{ij} = \eta_{zj} + \eta_{zi} = \underbrace{\mu_{\alpha\eta zj} + \zeta_{\alpha\eta zj}}_{\text{B part of } z_{ij}} + \underbrace{\zeta_{zi}}_{\substack{\text{W part} \\ \text{of } z_{ij}}}
\tag{24}
$$

$$
\begin{aligned}
y_{ij} &= \eta_{yj} + \eta_{yi} \\
&= \underbrace{\mu_{\alpha\eta yj} + \zeta_{\alpha\eta yj}}_{\text{random intercept}} + \underbrace{\beta_{y.x}}_{\substack{\text{fixed} \\ \text{slope for B} \\ \text{part of } x_{ij}}} \underbrace{(\mu_{\alpha\eta xj} + \zeta_{\alpha\eta xj})}_{\text{B part of } x_{ij}} + \underbrace{\beta_{y.z}}_{\substack{\text{fixed} \\ \text{slope for B} \\ \text{part of } z_{ij}}} \underbrace{(\mu_{\alpha\eta zj} + \zeta_{\alpha\eta zj})}_{\text{B part of } z_{ij}} + \underbrace{\beta_{y.xz}}_{\substack{\text{fixed intxn} \\ \text{of B parts of} \\ x_{ij} \text{ and } z_{ij}}} \underbrace{(\mu_{\alpha\eta xj} + \zeta_{\alpha\eta xj})}_{\text{B part of } x_{ij}} \underbrace{(\mu_{\alpha\eta zj} + \zeta_{\alpha\eta zj})}_{\text{B part of } z_{ij}} \\
&+ \underbrace{(\mu_{Byxj} + \zeta_{Byxj})}_{\substack{\text{random slope for} \\ \text{W part of } x_{ij}}} \underbrace{\zeta_{xi}}_{\substack{\text{W part} \\ \text{of } x_{ij}}} + \underbrace{(\mu_{Byzj} + \zeta_{Byzj})}_{\substack{\text{random slope for} \\ \text{W part of } z_{ij}}} \underbrace{\zeta_{zi}}_{\substack{\text{W part} \\ \text{of } z_{ij}}} + \underbrace{(\mu_{Byxizi} + \zeta_{Byxizi})}_{\substack{\text{random interaction of} \\ \text{W parts of } x_{ij} \text{ and } z_{ij}}} \underbrace{\zeta_{xi}}_{\substack{\text{W part} \\ \text{of } x_{ij}}} \underbrace{\zeta_{zi}}_{\substack{\text{W part} \\ \text{of } z_{ij}}} \\
&+ \underbrace{\mu_{Byxjzi}}_{\substack{\text{fixed intxn of} \\ \text{B part of } x_{ij} \\ \text{and W part of } z_{ij}}} \underbrace{\zeta_{zi}}_{\substack{\text{W part} \\ \text{of } z_{ij}}} \underbrace{(\mu_{\alpha\eta xj} + \zeta_{\alpha\eta xj})}_{\text{B part of } x_{ij}} + \underbrace{\mu_{Byxizj}}_{\substack{\text{fixed intxn of} \\ \text{W part of } x_{ij} \\ \text{and B part of } z_{ij}}} \underbrace{\zeta_{xi}}_{\substack{\text{W part} \\ \text{of } x_{ij}}} \underbrace{(\mu_{\alpha\eta zj} + \zeta_{\alpha\eta zj})}_{\text{B part of } z_{ij}} + \underbrace{\zeta_{yi}}_{\substack{\text{W residual} \\ \text{for } y_{ij}}}
\end{aligned}
\tag{25}
$$

Equation 25 contains all four possible interaction effects in $1 \times (1 \rightarrow 1)$ designs. In fact, terms in Equation 25 match those in Equation 6 with different symbols. Usually, not all of these effects will be of interest at once. Hypotheses B1, B2, C, and D can be tested by omitting rows and columns corresponding to the W parts of some variables in the model in Equations 16, 21, and 22.

To illustrate the LMS method, we again use an extension of SEM path diagrams. In Figure 1 Panel B, a B2 hypothesis (B interaction) is modeled by including the latent product $x_j \times z_j$ as a predictor of y_j, the random intercept for y_{ij}. For simplicity, we omitted the covariances between the interaction term and its components, i.e., x_j and z_j.

Advantages and Disadvantages of the RCP and LMS Methods

RCP and LMS allow testing interactions that are often overlooked and, when estimated using methods other than MSEM, may be subject to severe bias. Yet, RCP and LMS are not universally applicable; there are circumstances when one is preferred over the other.

First, researchers may want to model interaction effects as varying across clusters, such as the effect in Hypothesis A1 varying across teams. RCP cannot treat interactions as random slopes, but LMS can.[8] Second, RCP and LMS can be computation-intensive, requiring numerical integration and creative model specifications to achieve convergence. To ease estimation, it is possible to simultaneously apply RCP and LMS. For example, one could test Hypothesis B1 with RCP and B2 with LMS (as we later show). Third, with nonnormal predictors or moderators LMS yields biased *SE*s (Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Klein, Schermelleh-Engel, Moosbrugger, & Kelava, 2009) and biased point estimates (Coenders, Batista-Foguet, & Saris, 2008). Yet, for MSEM, the central limit theorem implies that cluster means become normal as a function of increasing cluster size, so cluster means may be normal even if L1 variables are not. As such, nonnormality may be less problematic for B parts of L1 variables, but the normality of any variable in a latent variable interaction is an issue. This said, simulations show that with moderate levels of nonnormality LMS is more efficient than competing approaches (see Cham et al., 2012). Fourth, RCP can model same-level interactions. However, we recommend LMS because RCP cannot estimate interactions among the W parts of L1 variables, and preliminary simulations showed good performance for LMS in the models we have described, but same-level RCP moderation was often biased and inefficient (see online supplemental materials).[9]

Finally, it is possible to use observed cluster means at the B level in lieu of latent cluster means for testing Hypotheses A2, A3, and B2 (Aguinis et al., 2013; Cronbach & Webb, 1975; Enders & Tofighi, 2007; Hofmann & Gavin, 1998; Ryu, 2015). In some cases this may be preferred, such as with sampling ratios near 1.0 (nearly all units within a cluster are sampled). In such cases, the L2 aggregate of the L1 variable is termed *formative*. More often, the observed cluster mean is considered a sample proxy for a population quantity, in which case it is termed *reflective* (for a discussion see Lüdtke et al., 2008). Using observed cluster means yields models that converge more quickly and are less prone to estimation errors. However, recent work (Preacher et al., 2011) shows that using observed cluster means in lieu of reflective latent aggregations leads to considerable bias in both the estimates of main effects and their *SE*s, even under the most favorable conditions of large cluster sizes and large ICCs. There is every reason to expect that using observed cluster means will result in poor performance in the multilevel moderation context in the reflective case. To support this assertion we ran a small simulation study that demonstrates the observed-means approach to have greater bias in both effects and *SE*s, and lower CI coverage than the MSEM approach with latent variables (see online supplemental materials).[10] However, because the W effects will be very similar using either observed or latent means in the B model, the observed means approach can be useful as a way to obtain starting values for MSEM in situations where convergence is challenging.

Simple Slopes Analysis for Multilevel Moderation Models

If statistically significant, an interaction effect is usually "plotted and probed" to examine a moderator's influence on the effect of a focal predictor by plotting regression lines at several values of the moderator (Aiken & West, 1991). Each regression line has a *simple intercept* and *simple slope* that are conditional on the moderator, often at its mean and $\pm$ 1 *SD*. Statistical significance of the intercepts and slopes is determined using *t* tests or with confidence intervals plotted as continuous functions of the moderator, forming confidence bands. These methods are used in single-level regression (Aiken & West, 1991), latent growth models (Curran, Bauer, & Willoughby, 2004), and multilevel models when L1 effects are not decomposed (Bauer & Curran, 2005; Curran, Bauer, & Willoughby, 2006; Preacher et al., 2006).

Similar procedures can be used for multilevel moderation (Raudenbush & Bryk, 1986), but this is less straightforward in MSEM because latent W and B parts of L1 variables are not observed. With latent variables assumed to be normal, conditional values based on estimated means and *SD*s of latent moderators can be chosen. For W latent parts, means of predictors and outcomes are zero because L1 variables are cluster-mean centered in MSEM. Otherwise, the same procedures may be used (Muthén, 2012). We now give an example of RCP and LMS.

Empirical Example: High School and Beyond

In this example we use a subsample of the 1982 High School and Beyond data with 7,185 students nested in 160 schools (see Raudenbush & Bryk, 2002). We show how to use theory to guide research design and make appropriate data analysis decisions by following *five simple steps*. Step 1: researchers must understand the interaction forms that are possible (see Table 1), and they should use substantive theory to guide their research questions, hypotheses, and model specifications. A great deal of literature suggests that poverty and school size interact in predicting achievement (e.g., Bickel & Howley, 2000; Cobbold, 2006). Specifically, the positive relationship between socioeconomic status and student achievement typically increases as school size increases, implying that smaller schools are better for low-SES students. We wish to determine whether this common finding can be attributed to moderation of the within-school effect of SES, the between-school effect, or both. In other words, we are interested in comparing and contrasting Hypotheses B1 and B2 with a $2 \times (1 \rightarrow 1)$ design as shown in Table 1.

Step 2: researchers must choose between a latent versus an observed-means approach. As shown below, the latter approach can bias point estimates and *SE*s, but this bias is reduced as cluster size, number of clusters, and ICCs increase. Without sampling error, the observed-means approach should be used, but in our example we want to correct for sampling error variance. Therefore, we specified latent means in a model with mathematics achievement (*math*) predicted by socioeconomic status (*ses*, an L1 predictor) and school enrollment (*size*, an L2 moderator, which is the number of students in a school divided by 1,000 to aid

[8] In Mplus, this involves regressing a phantom variable onto the latent product with a fixed slope, then regressing the outcome onto the phantom variable with a random slope. Example syntax is provided in the online appendix for this, along with syntax for all examples we discuss.

[9] An online appendix at http://quantpsy.org contains code for all of the models mentioned or illustrated in this article.

[10] We varied number of clusters (J = 50, 100, 200), cluster size (n_j = 5, 10, 20), and predictor ICC (.5 and .1). Results are available in the online appendix at http://quantpsy.org.

convergence). Because *ses* is an L1 predictor with B variance (ICC_{ses} = .27), it is plausible that (a) W standings on *ses* could exert an effect on W standings on *math*, and (b) latent B standings on *ses* could predict B standings on *math*. Moreover, either or both of these effects can be moderated by *size*. Thus, we specify a model testing Hypotheses B1 and B2 described earlier:

$$\begin{aligned} math_{ij} &= \beta_{0j} + \beta_{1j}ses_i + e_{ij} \\ \beta_{0j} &= \gamma_{00} + \gamma_{01}ses_{.j} + \gamma_{02}size_j + \gamma_{03}ses_{.j}size_j + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}size_j + u_{1j} \\ math_{ij} &= \gamma_{00} + \gamma_{10}ses_i + \gamma_{01}ses_{.j} + \gamma_{02}size_j \\ &\quad + \gamma_{11}ses_i size_j + \gamma_{03}ses_{.j}size_j + u_{0j} + u_{1j}ses_i + e_{ij} \end{aligned} \tag{28}$$

In this model, B1 is tested by examining $\hat{\gamma}_{11}$ and B2 is tested by examining $\hat{\gamma}_{03}$.

Step 3: the model is estimated with LMS and/or RCP. We discuss the weaknesses and strengths of these methods in a previous section and in the Discussion. To show the flexibility of our approach, we fit the model in two ways: (a) both the B1 and B2 hypotheses are tested via LMS, and (b) B1 is tested via RCP and B2 via LMS. Path diagrams are in Figure 2 and results are in Table 2. It is notable that when model estimation is difficult, researchers may consider alternative methods such as Bayes estimation or even an observed-means approach (despite its potential bias). In our case, results correspond closely across models, with a positive and significant cross-level interaction of *size* and within-school *ses*, such that the W part of *ses* has a positive effect on *math* that is stronger for larger schools and weaker for smaller schools. The interaction of the B part of *size* and B part of *ses* was negative but nonsignificant.

Step 4: using output from the LMS/LMS model, interactions are plotted to show the conditional effects. In our case this is a significant cross-level interaction, with the conditional effects of *ses* on *math* at five conditional values of *size*: the minimum, maximum, and 25th, 50th, and 75th percentiles (100, 555, 1,026, 1,436, and 2,713; see Figure 3).

Step 5: although usually unnecessary, for illustrative purposes we also estimate the model with an observed-means approach, with results in the last column of Table 2. Unsurprisingly, the results based on observed cluster means show clear downward biases in the point-estimate of the B interaction effect and in the *SE* of the cross-level interaction effect.

For comparison, the usual practice for testing interactions in $2 \times (1 \rightarrow 1)$ designs is to estimate a model as:

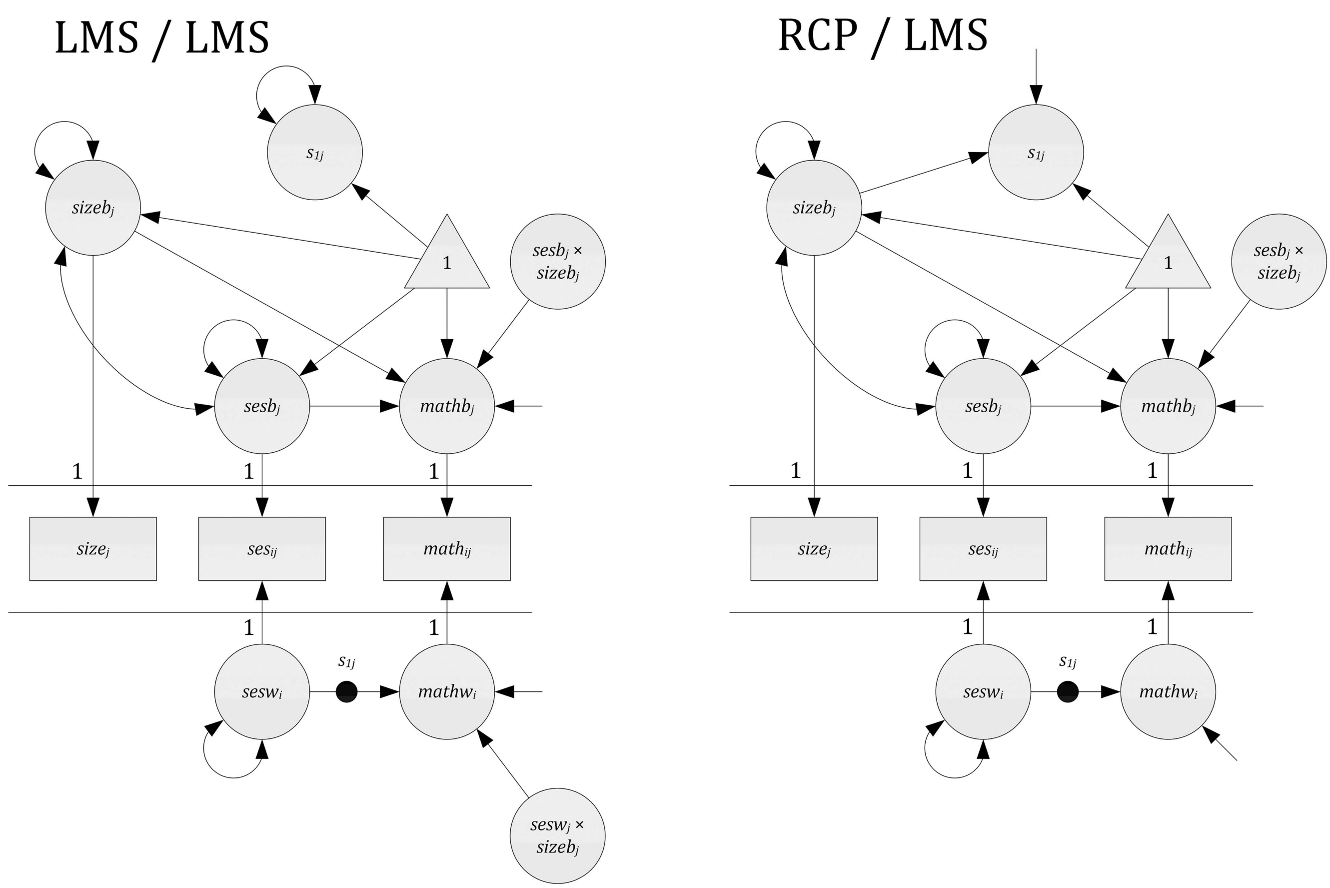

Figure 2. Path diagrams representing two different approaches to simultaneously testing hypotheses B1 and B2 with the High School and Beyond data. Covariances involving the random slope and the product terms are estimated, but omitted from the diagram for clarity. RCP = random coefficient prediction; LMS = latent moderated structural equations.

Table 2
Results From Analyses of High School and Beyond Data: Level-Specific Effects

Parameter	Method (Within/Between)		
	LMS/LMS	RCP/LMS	Traditional (MLR)
$ses_B \rightarrow math_B$	**7.070** (.745)	**7.070** (.745)	**6.215** (.606)
$size_B \rightarrow math_B$	−.079 (.256)	−.077 (.256)	−.073 (.253)
$ses_B \times size_B \rightarrow math_B$	−.515 (.588)	−.516 (.588)	−.296 (.504)
$ses_W \rightarrow math_W$	**1.607** (.255)	**1.608** (.255)	1.585 (.249)
$ses_W \times size_B \rightarrow math_W$	**.584** (.221)	**.583** (.221)	**.554** (.010)
mean(ses_B)	−.009 (.033)	−.009 (.033)	
mean($size_B$)	**1.098** (.050)	**1.098** (.050)	
intercept($math_B$)	**12.778** (.328)	**12.775** (.328)	**12.764** (.326)
var(ses_B)	**.151** (.017)	**.151** (.017)	
var($size_B$)	**.394** (.037)	**.394** (.037)	
var(u_{0j})	**2.062** (.440)	**2.063** (.441)	**2.641** (.465)
var(u_{1j})	**.595** (.262)	**.599** (.263)	**.588** (.260)
cov(u_{0j},u_{1j})	−.267 (.246)	−.264 (.247)	−.241 (.240)
cov(u_{1j},ses_B)	.069 (.061)	.069 (.061)	
cov($size_B$,ses_B)	−.032 (.023)	−.032 (.023)	
var(ses_W)	**.436** (.010)	**.436** (.010)	
var(e_{ij})	**36.633** (.720)	**36.642** (.720)	**36.701** (.721)

Note. Numbers in parentheses are standard errors. Numbers in boldface are significant at α = .05. LMS = latent moderated structural equations; RCP = random coefficient prediction; Traditional = unconflated multilevel model using observed means; MLR = maximum likelihood estimation with robust standard errors.

$$
\begin{aligned}
math_{ij} &= \beta_{0j} + \beta_{1j} ses_{ij} + e_{ij} \\
\beta_{0j} &= \gamma_{00} + \gamma_{01} size_j + u_{0j} \\
\beta_{1j} &= \gamma_{10} + \gamma_{11} size_j + u_{1j} \\
math_{ij} &= \gamma_{00} + \gamma_{10} ses_{ij} + \gamma_{01} size_j \\
&\quad + \gamma_{11} ses_{ij} size_j + u_{0j} + u_{1j} ses_{ij} + e_{ij}
\end{aligned}
\tag{29}
$$

which conflates the possible B and cross-level interaction effects of *ses* and *size*, yielding the results in Table 3. The single interaction effect, while still significant, is a weighted average of the B1 and B2 interaction effects estimated in Equation 28. Alternatively, a single conflated interaction effect does not show that the magnitude and significance of that effect reflects moderation of the W effect (rather than the B effect) of *ses* on *math* by the L2 variable *size*—showing why we call this approach "conflated."

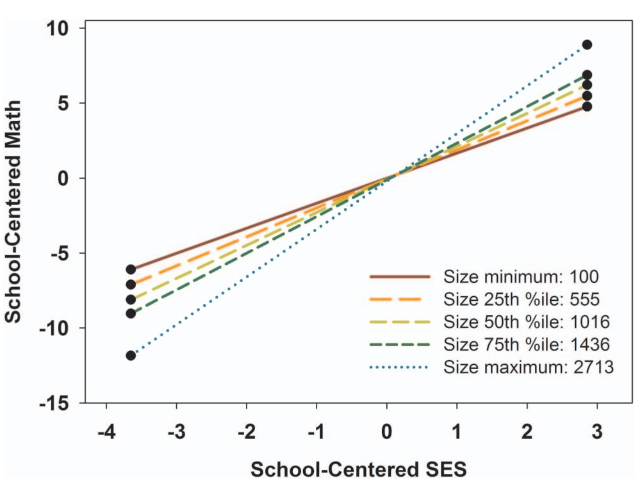

Figure 3. Cross-level interaction of *ses* and *size*. See the online article for the color version of this figure.

Discussion

This article treats RCP and LMS approaches to multilevel moderation in MSEM. Although our focus is methodological, our work has implications for theory development, which is done within the perceived limits of available methods. Current approaches to multilevel moderation are often tied to a single-level logic, making it difficult to tie multilevel theory to multilevel methods and produce theories and hypotheses that reflect the complexity and specificity of MSEM. Knowing the types of moderation that are possible (e.g., interactions involving the B or W parts of L1 variables) enables researchers to clearly articulate and test moderation in multilevel research.

In turn, there is a disturbing implication of our work: Because many studies of multilevel moderation use methods developed for single-level data, many moderation effects estimated with multilevel models may be "uninterpretable blends" of interactions involving W and B parts (Cronbach, 1976, p. 9.20). It is possible that many legitimate effects have escaped notice because L1 variables

Table 3
Results From Analyses of High School and Beyond Data: Conflated Multilevel Model

Parameter	Estimate (*SE*)
$ses \rightarrow math$	**1.852** (.236)
$size \rightarrow math$	−.261 (.308)
$ses \times size \rightarrow math$	**.492** (.191)
intercept($math$)	**12.973** (.425)
var(u_{0j})	**4.804** (.757)
var(u_{1j})	.314 (.221)
cov(u_{0j},u_{1j})	−.102 (.325)
var(e_{ij})	**36.829** (.725)

Note. Numbers in parentheses are standard errors. Numbers in boldface are significant at α = .05. The random coefficient prediction (RCP) method was used.

and their effects were not separated into B and W parts. Researchers who fail to locate a significant interaction may have simply failed to separate cross-level interactions into statistically and practically significant W and B parts of opposite sign (similar to problems in multilevel mediation analysis; see Preacher et al., 2011, 2010). Alternatively, many studies that found cross-level interactions actually may have found a B interaction conflated with a small or absent cross-level interaction. The point is that prior research may have missed opportunities to uncover theoretically interesting interactions.

Our approach avoids such difficulties by clearly articulating multilevel moderation effects, estimating these effects, and interpreting them correctly. This is key because the implications of scientific findings involve policy and interventions which, targeted at the wrong level or unit of analysis, are less likely to have their intended effects. Although existing work emphasizes such thinking for multilevel direct and indirect effects (e.g., Preacher et al., 2010), decomposing multilevel moderation effects may be new to most researchers. We now treat several issues relevant to interpreting interactions as well as limitations and extensions of our work.

Interpreting Multilevel Moderation Effects

Moving from conflated effects to decomposed effects. Multilevel researchers often adopt an MLM strategy of starting with conflated W and B effects, and only later consider centering L1 predictors and including cluster means as predictors (a "*conflated-first*" strategy). We argue for the opposite strategy—starting with decomposed effects and theory regarding their potential equality (a "*decomposed-first*" strategy). MSEM proceeds in the latter way; the W and B parts of any L1 variable are automatically decomposed and researchers are motivated to theorize why they would explicitly constrain any effects to equality in order to obtain results comparable with those from traditional MLM with no centering or grand-mean centering.

This strategy has many benefits, such as bringing multilevel theory to the fore, making different effects across levels obvious, and allowing equality constraints on B and W effects so that models with decomposed L1 variables are similar to traditional MLM. A decomposed-first strategy using MSEM also allows the B effect of an L1 variable to be an unbiased estimate of the latent cluster-level component's effect. Further, it is easy to test invariance in W and B effects because a model with them constrained to equality is parametrically nested in one with them freely estimated. By separating and freely estimating W and B effects, MSEM automatically produces an alternative model to which more constrained models can be compared. On the other hand, when the B part of a model is not of interest, MSEM makes it easy to allow all B variables to freely correlate, allowing a focus on only a model's W part and its statistical fit. Level-specific main effects and interactions may not always be of interest to researchers, and in this case effects need not be decomposed. Yet, we argue that the potential for theoretically interesting, level-specific effects is overlooked by many researchers using the conflated-first strategy, and it deserves more attention in both the methodological and applied literatures.

Interpreting cross-level interactions. The benefits of a decomposed-first strategy are shown in the case of cross-level interactions, which are usually said to occur if L2 variables moderate the slope of L1 predictors. We suggest that L1 variables and effects should be decomposed into level-specific parts, and theory may suggest that either or both of the B and W effects of L1 variables are moderated, indicating how cross-level interactions are specified and interpreted (Cronbach & Webb, 1975; Cronbach & Snow, 1977, Chapter 4; Enders & Tofighi, 2007).

Consider the typical cross-level interaction in Equation 7, repeated here as:

$$y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + \gamma_{01}z_j + \gamma_{11}x_{ij}z_j + u_{0j} + u_{1j}x_{ij} + \varepsilon_{ij} \quad (30)$$

Here, the simple slope of y_{ij} regressed on x_{ij} is $(\gamma_{10} + \gamma_{11}z_j)$, and only γ_{11} is an interaction effect. Yet, because x_{ij} has W and B parts, so does its slope, and thus so does its simple slope:

$$\begin{aligned} y_{ij} = \gamma_{00} &+ \gamma_{01}x_{.j} + \gamma_{02}z_j + \gamma_{03}x_{.j}z_j + \gamma_{10}x_i \\ &+ \gamma_{11}x_iz_j + u_{0j} + u_{1j}x_i + \varepsilon_{ij} \end{aligned} \quad (31)$$

In Equation 31, γ_{03} is a strictly B interaction because both $x_{.j}$ and z_j are B variables. Yet, γ_{11} is a truly cross-level interaction. Although the product x_iz_j varies at two levels, it does not explain any W variance in y_{ij} beyond that explained by x_i (within a cluster, x_i and x_iz_j correlate at 1). Thus, γ_{11} is the degree to which the slope relating two strictly W variables varies as a function of a strictly B variable. The traditional cross-level interaction in Equation 30 is only partially cross-level because it includes elements of a truly cross-level interaction and a strictly B interaction—emphasized by Cronbach and colleagues decades ago but mostly forgotten today. Researchers can test the equality of these interactions by constraining γ_{03} and γ_{11} to equality and testing fit, but these interactions are incommensurable because one is a change in a B effect (γ_{03}) and the other is a change in a W effect (γ_{11}). Thus, the decomposed-first strategy for cross-level interactions shows how the W and B effects of x_{ij} may be differently moderated.

Implementation Issues

When applying our MSEM approach there are various concerns to address. In order to develop and build models for multilevel moderation in MSEM, the usual sequence of model building in SEM and moderation testing still applies. In some situations researchers may begin with a "main effects only" model and then proceed to include interactions. If theory suggests moderation, researchers may begin with a model involving interactions. Precisely which interactions to include should depend on theory. Just as with multilevel mediation in MSEM (Preacher et al., 2010), researchers should check assumptions associated with their data and models, including any measurement properties of L1 scales at multiple levels of analysis (Geldhof, Preacher, & Zyphur, 2014). Also, work on model-building in single-level SEM has yet to be adequately generalized to the multilevel case, and future work should address this topic. Here, we discuss some new concerns that our approach creates in choosing between RCP and LMS, and how to tackle centering predictors, a central topic in MLM and moderation literature.

Choosing between RCP and LMS. As we note, RCP and LMS are not equally applicable to all types of multilevel moderation. Strictly W interactions cannot be estimated with RCP, but LMS allows this. Alternatively, RCP is suited for cross-level interactions with random slopes because it was designed for this purpose and researchers are familiar with it. Further, as we noted, RCP can encounter bias, inefficiency, and convergence issues if researchers use it for B-level interactions. As such, we recommend RCP for cross-level interactions, where it performs well. On the other hand, the performance of LMS (in terms of bias, coverage, and Type I error control) declines under nonnormality for the interacting predictors (Cham et al., 2012). Beyond these concerns, the choice between RCP and LMS can be made to limit estimation difficulty and computation time. Especially in complex models that allow mixing the use of RCP and LMS for cross- and same-level interactions, respectively, researchers may have to experiment with their models to examine what works best for their specific data.

In the future, computation difficulties will be reduced with innovations in estimation, such as Monte Carlo integration or Bayes estimation, the latter of which is possible for RCP in Mplus but is not yet implemented for LMS. As always, software developments will allow a wider range of possible models and estimators to reduce computation issues. Regardless of the forms such advances take, the decomposition-based MSEM approach outlined here will apply.

Centering. A key question in multilevel modeling is whether to group- or grand-mean center L1 variables (e.g., Enders & Tofighi, 2007; Raudenbush & Bryk, 2002). Literature on moderation often advocates grand-mean centering predictors and moderators (e.g., Aiken & West, 1991; Cronbach & Snow, 1977). Multilevel moderation tests are at the intersection of these two literatures, so centering decisions have important implications.

In MLM, group-mean centering L1 predictors eliminates their B parts, leaving only their W parts, but MSEM automatically decomposes L1 variables, so this is a nonissue. In MSEM, grand-mean centering will give similar results to raw scores *except in one case*. When a random slope for a W effect is regressed on a predictor (either a L1 variable's B part or an observed L2 variable), the intercept of the random slope is adjusted by the mean of the moderator. Thus, the random slope's intercept is the W effect with the moderator set to 0. Centering the moderator changes the random slope's intercept, and thus the interpretation of the L1 predictor's conditional main effect. Grand-mean centering or using raw scores does not affect results, but the potential to misinterpret the mean of a random slope as a W effect should be kept in mind.

Limitations

A major advantage of MSEM is that it allows treating the B part of L1 variables as latent, removing bias caused by using cluster means at L2 (Lüdtke et al., 2008; Marsh et al., 2009). Also, with MSEM, all types of multilevel moderation can be embedded in more complex models, such as mixture models (e.g., Muthén & Asparouhov, 2009; Sterba & Bauer, 2014), mediation models (e.g., Preacher et al., 2010), and models with multiple-indicator latent variables. Yet, the potential for such complex models hints at a limitation of MSEM: convergence issues and computation times. Advances in estimation and computing power will solve some of these, but the problems are difficult to overcome because the complexity of researchers' models and designs routinely push the boundaries of what is practical. There are no easy solutions to these problems, and researchers often will have to adjudicate among, and experiment with, the models they would like to estimate and those that are practically feasible (as in Step 2 above).

Potential Extensions

Our work serves as a basis for testing *conditional indirect effects* with multilevel data, or multilevel *moderated mediation*, which exists when indirect effects are functions of moderators (Edwards & Lambert, 2007; Hayes, 2013; Muller, Judd, & Yzerbyt, 2005; Preacher, Rucker, & Hayes, 2007). Such effects in multilevel designs have yet to receive serious attention, with only two articles addressing the topic. Bauer, Preacher, and Gil (2006) present a method for examining a special case: when a conflated indirect effect is moderated by an L2 moderator. Ryu (2015) considers MSEM for moderated mediation, but uses Muthén's (1990) approach, requiring complete data and balanced cluster sizes with no random slopes or cross-level interactions. Our approach does not have these limitations and allows the W and B effects of L1 variables to be conditional on moderators assessed at any level. If such moderated effects or even random slopes participate in longer causal chains (i.e., indirect effects), then level-specific or cross-level moderated mediation exists. Moderation of multilevel mediation combines our logic with that of Preacher et al. (2011, 2010), which in turn is a generalization of work by Edwards and Lambert (2007) and Preacher et al. (2007) to MSEM in the B, W, and cross-level cases.

MSEM also allows higher-order interactions and additional levels of analysis. With LMS, higher-order (e.g., 3-way) interactions require multiplying the result of a latent interaction by a moderator, creating a higher-order term that can be used to estimate interactions. Such models, as well as those we treat here, can be easily extended to three-level data. We do not discuss these models here, but there is no theoretical impediment to specifying them. As above, researchers simply need to focus on appropriately decomposing observed variables into level-specific parts.[11]

In sum, given the generality of MSEM there are many other potential variations on the models we treat. We cannot discuss all possibilities, but we do emphasize a feature of MSEM that overcomes a key shortcoming of MLM: traditional MLM limits researchers to investigating outcomes at L1, but MSEM allows L1 or L2 variables as outcomes (Muthén & Asparouhov, 2009; Preacher et al., 2010). Indeed, L2 outcomes with interaction effects are simple to specify in MSEM by treating an L2 outcome as an element of the y_{ij} vector (in Equation 12) for which only the B terms would be specified (in Equation 14). This allows tests of moderation in designs such as $1 \times (1 \rightarrow 2)$, $1 \times (2 \rightarrow 2)$, and $2 \times$

[11] Mplus currently tests moderation in three-level models using only RCP and the Bayes estimator, implying that interactions among W latent components of L1 variables are currently infeasible in three-level designs.

(1 → 2).[12] With an L2 outcome, all effects are necessarily B effects, restricting moderation types to those involving observed or latent B components.

Conclusion

In this article we have (a) clarified how current methods for assessing multilevel moderation can result in biased or misleading results; (b) provided methods that yield appropriate and interpretable results for different types of multilevel moderation; and (c) encouraged the implementation and use of stable, efficient estimators capable of fitting such models. Our hope is that our contributions will spur further research in these areas.

[12] 2 × (2 → 2) designs ordinarily would be treated using single-level analysis, unless the model includes covariates measured at a lower or higher level.

References

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management, 39,* 1490–1528. http://dx.doi.org/10.1177/0149206313478188

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

Asparouhov, T., & Muthén, B. (2006). *Constructing covariates in multilevel regression*. Mplus Web Notes: No. 11. Retrieved from http://www.statmodel.com

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research, 40,* 373–400. http://dx.doi.org/10.1207/s15327906mbr4003_5

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods, 11,* 142–163. http://dx.doi.org/10.1037/1082-989X.11.2.142

Bickel, R., & Howley, C. (2000). The influence of scale on school performance: A multi-level extension of the Matthew principle. *Education Policy Analysis Archives, 8,* 22–32. http://dx.doi.org/10.14507/epaa.v8n22.2000

Cham, H., West, S. G., Ma, Y., & Aiken, L. S. (2012). Estimating latent variable interactions with nonnormal observed data: A comparison of four approaches. *Multivariate Behavioral Research, 47,* 840–876. http://dx.doi.org/10.1080/00273171.2012.732901

Chou, C., Bentler, P. M., & Pentz, M. A. (2000). A two-stage approach to multilevel structural equation models: Application to longitudinal data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches and specific examples* (pp. 33–50). Mahwah, NJ: Erlbaum.

Cobbold, T. (2006). *The interaction of school size and socio-economic status on student performance*. Unpublished manuscript. Retrieved from http://www.saveourschools.com.au/

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

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

Cronbach, L. J. (1976). *Research on classrooms and schools: Formulation of questions, design, and analysis*. Stanford, CA: Stanford University Evaluation Consortium.

Cronbach, L. J., & Snow, R. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York, NY: Irvington Publishers.

Cronbach, L. J., & Webb, N. (1975). Between-class and within-class effects in a reported aptitude × treatment interaction: Reanalysis of a study by G. L. Anderson. *Journal of Educational Psychology, 67,* 717–724. http://dx.doi.org/10.1037/0022-0663.67.6.717

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology, 62,* 583–619. http://dx.doi.org/10.1146/annurev.psych.093008.100356

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2004). Testing main effects and interactions in latent curve analysis. *Psychological Methods, 9,* 220–237. http://dx.doi.org/10.1037/1082-989X.9.2.220

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2006). Testing and probing interactions in hierarchical linear growth models. In C. S. Bergeman & S. M. Boker (Eds.), *The Notre Dame series on quantitative methodology, Vol. 1*: *Methodological issues in aging research* (pp. 99–129). Mahwah, NJ: Erlbaum.

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods, 12,* 1–22. http://dx.doi.org/10.1037/1082-989X.12.1.1

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12,* 121–138. http://dx.doi.org/10.1037/1082-989X.12.2.121

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods, 19,* 72–91. http://dx.doi.org/10.1037/a0032138

Goldstein, H., & McDonald, R. P. (1988). A general model for the analysis of multilevel data. *Psychometrika, 53,* 455–467. http://dx.doi.org/10.1007/BF02294400

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46,* 1251–1271. http://dx.doi.org/10.2307/1913827

Hayes, A. F. (2013). *An introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis*. Hoboken, NJ: Wiley.

Hildreth, C., & Houck, J. P. (1968). Some estimates for a linear model with random coefficients. *Journal of the American Statistical Association, 63,* 584–595.

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management, 24,* 623–641. http://dx.doi.org/10.1177/014920639802400504

Jaccard, J. J., & Turrisi, R. (2003). *Interaction effects in multiple regression* (2nd ed.). Thousand Oaks, CA: Sage.

Jedidi, K., & Ansari, A. (2001). Bayesian structural equation models for multilevel data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 129–157). New York, NY: Erlbaum.

Johnston, J. (1984). *Econometric methods* (3rd ed.). New York, NY: McGraw-Hill.

Jöreskog, K. G., & Yang, F. (1996). Nonlinear structural equation models: The Kenny-Judd model with interaction effects. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 57–88). Mahwah, NJ: Erlbaum.

Kelava, A., Moosbrugger, H., Dimitruk, P., & Schermelleh-Engel, K. (2008). Multicollinearity and missing constraints: A comparison of three approaches for the analysis of latent nonlinear effects. *Methodology, 4,* 51–66. http://dx.doi.org/10.1027/1614-2241.4.2.51

Kelava, A., Werner, C., Schermelleh-Engel, K., Moosbrugger, H., Zapf, D., Ma, Y., . . . West, S. G. (2011). Advanced nonlinear structural equation modeling: Theoretical properties and empirical application of the LMS and QML estimators. *Structural Equation Modeling, 18,* 465–491. http://dx.doi.org/10.1080/10705511.2011.582408

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin, 96,* 201–210. http://dx.doi.org/10.1037/0033-2909.96.1.201

Kidwell, R. E., Jr., Mossholder, K. W., & Bennett, N. (1997). Cohesiveness and organizational citizenship behavior: A multilevel analysis using work groups and individuals. *Journal of Management, 23,* 775–793. http://dx.doi.org/10.1177/014920639702300605

Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika, 65,* 457–474. http://dx.doi.org/10.1007/BF02296338

Klein, A. G., & Muthén, B. (2007). Quasi-maximum likelihood of structural equation modeling with multiple interaction and quadratic effects. *Multivariate Behavioral Research, 42,* 647–673. http://dx.doi.org/10.1080/00273170701710205

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

Kreft, I. G. G., de Leeuw, J., & Aiken, L. S. (1995). The effects of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research, 30,* 1–21. http://dx.doi.org/10.1207/s15327906mbr3001_1

Leite, W. L., & Zuo, Y. (2011). Modeling latent interactions at level 2 in multilevel structural equation models: An evaluation of mean-centered and residual-centered unconstrained approaches. *Structural Equation Modeling, 18,* 449–464. http://dx.doi.org/10.1080/10705511.2011.582400

Little, T. D., Bovaird, J. A., & Widaman, K. F. (2006). On the merits of orthogonalizing powered and product terms: Implications for modeling interactions among latent variables. *Structural Equation Modeling, 13,* 497–519. http://dx.doi.org/10.1207/s15328007sem1304_1

Lüdtke, O., Marsh, H. W., Robitzsch, A., & Trautwein, U. (2011). A 2 × 2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error correction models. *Psychological Methods, 16,* 444–467. http://dx.doi.org/10.1037/a0024376

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods, 13,* 203–229. http://dx.doi.org/10.1037/a0012869

Mancl, L. A., Leroux, B. G., & DeRouen, T. A. (2000). Between-subject and within-subject statistical information in dental research. *Journal of Dental Research, 79,* 1778–1781. http://dx.doi.org/10.1177/00220345000790100801

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44,* 764–802. http://dx.doi.org/10.1080/00273170903333665

Marsh, H. W., Wen, Z., & Hau, K.-T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods, 9,* 275–300. http://dx.doi.org/10.1037/1082-989X.9.3.275

Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology, 97,* 951–966. http://dx.doi.org/10.1037/a0028380

McDonald, R. P. (1993). A general model for two-level data with responses missing at random. *Psychometrika, 58,* 575–585. http://dx.doi.org/10.1007/BF02294828

McDonald, R. P. (1994). The bilevel reticular action model for path analysis with latent variables. *Sociological Methods & Research, 22,* 399–413. http://dx.doi.org/10.1177/0049124194022003007

McDonald, R. P., & Goldstein, H. I. (1989). Balanced versus unbalanced designs for linear structural relations in two-level data. *British Journal of Mathematical and Statistical Psychology, 42,* 215–232. http://dx.doi.org/10.1111/j.2044-8317.1989.tb00911.x

Moosbrugger, H., Schermelleh-Engel, K., Kelava, A., & Klein, A. G. (2009). Testing multiple nonlinear effects in structural equation modeling: A comparison of alternative estimation approaches. In T. Teo & M. S. Khine (Eds.), *Structural equation modeling in educational research: Concepts and applications* (pp. 103–135). Rotterdam: Sense Publishers.

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89,* 852–863. http://dx.doi.org/10.1037/0022-3514.89.6.852

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54,* 557–585. http://dx.doi.org/10.1007/BF02296397

Muthén, B. O. (1990). *Mean and covariance structure analysis of hierarchical data*. Los Angeles, CA: UCLA Statistics series, No. 62.

Muthén, B. O. (1991). Multilevel factor analysis of class and student achievement components. *Journal of Educational Measurement, 28,* 338–354. http://dx.doi.org/10.1111/j.1745-3984.1991.tb00363.x

Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research, 22,* 376–398. http://dx.doi.org/10.1177/0049124194022003006

Muthén, B. O. (2011). *Applications of causally defined direct and indirect effects in mediation analysis using SEM in Mplus*. Unpublished manuscript. Retrieved from http://www.statmodel.com/download/causalmediation.pdf

Muthén, B. O. (2012). *Latent variable interactions*. Unpublished manuscript. Retrieved. http://statmodel2.com/download/Latent%20variable%20interactions.pdf

Muthén, B. O., & Asparouhov, T. (2009). Growth mixture modeling: Analysis with non-Gaussian random effects. In G. Fitzmaurice, M. Davidian, G. Verbeke, & G. Molenberghs (Eds.), *Longitudinal data analysis* (pp. 143–165). Boca Raton, FL: Chapman & Hall/CRC.

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology, 25,* 267–316. http://dx.doi.org/10.2307/271070

Muthén, L. K., & Muthén, B. O. (1998–2014). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Neuhaus, J. M., & Kalbfleisch, J. D. (1998). Between- and within-cluster covariate effects in the analysis of clustered data. *Biometrics, 54,* 638–645. http://dx.doi.org/10.2307/3109770

Neuhaus, J., & McCulloch, C. (2006). Separating between- and within-cluster covariate effects by using conditional and partitioning methods. *Journal of the Royal Statistical Society Series B. Methodological, 68,* 859–872. http://dx.doi.org/10.1111/j.1467-9868.2006.00570.x

Preacher, K. J. (2011). Multilevel SEM strategies for evaluating mediation in three-level data. *Multivariate Behavioral Research, 46,* 691–731. http://dx.doi.org/10.1080/00273171.2011.589280

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31,* 437–448. http://dx.doi.org/10.3102/10769986031004437

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42,* 185–227. http://dx.doi.org/10.1080/00273170701341316

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2011). Alternative methods for assessing mediation in multilevel data: The advantages of multilevel SEM. *Structural Equation Modeling, 18,* 161–182. http://dx.doi.org/10.1080/10705511.2011.557329

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15,* 209–233. http://dx.doi.org/10.1037/a0020141

Raudenbush, S. W. (1989a). "Centering" predictors in multilevel analysis: Choices and consequences. *Multilevel Modelling Newsletter, 1,* 10–12.

Raudenbush, S. W. (1989b). A response to Longford and Plewis. *Multilevel Modelling Newsletter, 1,* 8–11.

Raudenbush, S. W., & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education, 59,* 1–17. http://dx.doi.org/10.2307/2112482

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

Raudenbush, S. W., & Sampson, R. (1999). Assessing direct and indirect effects in multilevel designs with latent variables. *Sociological Methods & Research, 28,* 123–153. http://dx.doi.org/10.1177/0049124199028002001

Ryu, E. (2015). The role of centering for interaction of level 1 variables in multilevel structural equation models. *Structural Equation Modeling, 22,* 617–630.

Schermelleh-Engel, K., Klein, A., & Moosbrugger, H. (1998). Estimating nonlinear effects using a latent moderated structural equations approach. In R. E. Schumacker & G. A. Marcoulides (Eds.), *Interaction and nonlinear effects in structural equation modeling* (pp. 203–238). Mahwah, NJ: Erlbaum.

Schmidt, W. H. (1969). *Covariance structure analysis of the multivariate random effects model* (Unpublished doctoral dissertation). University of Chicago, Chicago, IL.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal and structural equation models*. Boca Raton, FL: Chapman & Hall/CRC. http://dx.doi.org/10.1201/9780203489437

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Thousand Oaks, CA: Sage.

Sterba, S. K., & Bauer, D. J. (2014). Predictions of individual change recovered with latent class or random coefficient growth models. *Structural Equation Modeling, 21,* 342–360. http://dx.doi.org/10.1080/10705511.2014.915189

Van Yperen, N. W., & Snijders, T. A. B. (2000). A multilevel analysis of the demands—Control model: Is stress at work determined by factors at the group level or the individual level? *Journal of Occupational Health Psychology, 5,* 182–190. http://dx.doi.org/10.1037/1076-8998.5.1.182

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Yuan, K.-H., Cheng, Y., & Maxwell, S. (2014). Moderation analysis using a two-level regression model. *Psychometrika, 79,* 701–732. http://dx.doi.org/10.1007/s11336-013-9357-x

Received January 11, 2014
Revision received May 7, 2015
Accepted June 16, 2015 ■