Published in final edited form as:

Br J Math Stat Psychol. 2013 May ; 66(2): 290–307. doi:10.1111/j.2044-8317.2012.02051.x.

Multilevel mediation analysis: The effects of omitted variables in the 1–1–1 model

Davood Tofighi^{1,*}, Stephen G. West², and David P. MacKinnon²

¹Georgia Institute of Technology, USA

²Arizona State University, USA

Abstract

Multilevel mediation analysis examines the indirect effect of an independent variable on an outcome achieved by targeting and changing an intervening variable in clustered data. We study analytically and through simulation the effects of an omitted variable at level 2 on a 1–1–1 mediation model for a randomized experiment conducted within clusters in which the treatment, mediator, and outcome are all measured at level 1. When the residuals in the equations for the mediator and the outcome variables are fully orthogonal, the two methods of calculating the indirect effect (ab , $c - c'$) are equivalent at the between- and within-cluster levels. Omitting a variable at level 2 changes the interpretation of the indirect effect and will induce correlations between the random intercepts or random slopes. The equality of within-cluster ab and $c - c'$ no longer holds. Correlation between random slopes implies that the within-cluster indirect effect is conditional, interpretable at the grand mean level of the omitted variable.

1. Introduction

Mediation analysis is a statistical approach used to understand how an independent variable produces an indirect effect on an outcome through an intervening variable (mediator). To cite two examples, a preventive intervention (programme booklet versus no treatment control) might be hypothesized to affect risk factors (improved substance use refusal skills). Changes in this risk factor, in turn, are expected to decrease negative outcomes (substance use). A diet programme might be hypothesized to reduce food intake, which, in turn, is hypothesized to reduce the participant's body mass index. An indirect (mediated) effect is defined conceptually as the effect of the programme on the outcome that is transmitted through the mediator. The basic mediation model is depicted in Figure 1.

Building on Baron and Kenny's (1986) seminal work,¹ mediation models for analysing single-level data (i.e., a data structure without clustering) have undergone extensive

*Correspondence should be addressed to Davood Tofighi, School of Psychology, J. S. Coon Building, Georgia Institute of Technology, 654 Cherry Street, Atlanta, GA 30332-0170, USA (dtofighi@psych.gatech.edu).

Supporting Information

The following supporting information may be found in the online edition of this article:

Extended analytical proofs

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

development (MacKinnon, 2008; MacKinnon, Fairchild, & Fritz, 2007). More recently mediation analysis has been extended to the multilevel context (e.g., Bauer, Preacher, & Gil, 2006; Kenny, Korchmaros, & Bolger, 2003; Krull & MacKinnon, 1999). In the 1–1–1 mediation model in which the treatment, mediator, and outcome are all measured at level 1, confusion has resulted because the two methods of calculating the indirect effect (ab , $c - c'$, described below) are not equal as they are in the single-level case. We show below that when the residuals at both level 1 and level 2 are fully orthogonal the two methods of calculating the indirect effect are equivalent. We then explore analytically the effects of permitting a correlation between the random effects. Importantly, when a correlation is permitted between the random slopes the interpretation of the indirect effect is altered. We offer practical advice on how to address and interpret a potential cause of correlation between the random effects in the 1–1–1 mediational model.

2. The single-level mediation model

Consider a single-level randomized controlled trial with two groups in which the independent variable (intervention programme versus control) is hypothesized to change a mediator (e.g., M = substance use refusal skills) which, in turn, changes an outcome variable (e.g., Y = frequency of substance use). Three equations used to assess quantities in the single-mediator model are as follows (MacKinnon, 2008):

$$Y_i = d_1 + cX_i + \varepsilon_{1i}, \quad (1)$$

$$M_i = d_2 + aX_i + \varepsilon_{2i}, \quad (2)$$

$$Y_i = d_3 + c'X_i + bM_i + \varepsilon_{3i}, \quad (3)$$

where Y_i is the outcome variable measured on individual i , X_i is an indicator variable that represents whether the i th person received the intervention (1 = programme, 0 = control), and M_i is the mediator. The coefficient c in equation (1) represents the total effect of the prevention programme on substance use. The coefficient c' in equation (3) represents the direct effect of the prevention programme on substance use, controlling for the participants' refusal skills. The direct effect indicates the part of the total programme effect not accounted for by the mediator. The coefficient b describes the effect of refusal skills on substance use, controlling for the programme. The coefficient a in equation (2) represents the degree to which the intervention increased refusal skills relative to the control group. The residuals of equations (2) and (3) are assumed to have full orthogonality (discussed in more detail below), which implies that ε_{2i} and ε_{3i} have zero correlation (Bollen, 1989; McDonald, 1997). In addition, for purposes of estimation and interpretation using maximum likelihood, the residuals ε_{2i} and ε_{3i} are each assumed to have a normal distribution. These assumptions imply that the two error terms are independently and normally distributed across individuals. The magnitude of the effect of the prevention programme on decreasing substance use

¹As of January 2012, Baron and Kenny (1986) had over 30,000 citations according to Google Scholar illustrating the exponential increase in the use of mediation analysis in basic and applied research.

mediated by the individuals' refusal skills is represented by the indirect effect, the product of two coefficients, ab , with its estimate denoted by $\hat{ab}$. Another equivalent measure of the indirect effect in the single-level case is $c - c'$ (MacKinnon, Warsi, & Dwyer, 1995).

3. The basic multilevel mediation model

Many interventions occur in group settings (clusters, contexts) such as schools or community groups (e.g., Botvin & Eng, 1980; Ellickson, McCaffrey, Ghosh-Dastidar, & Longshore, 2003), even when the programmes themselves are delivered separately to each individual within the cluster. In group settings the individuals within a cluster tend to be more similar than individuals selected from different clusters, thereby violating the statistical assumption of independence of units. Individuals within a group may produce more homogeneous responses because of factors such as within-cluster participant interaction, influences from shared environmental factors, and similarity of group members. Such clustered data are referred to as *multilevel* data; data collection can potentially occur at multiple levels. In a two-level design, the data on individuals are referred to as level-1 data, whereas the data collected on the clusters (groups) are referred to as level-2 data. Statistical techniques such as multilevel analysis that account for clustering are needed (Hawkins, Brown, Oesterle, Arthur, Abbott, & Catalano, 2008) or clustering may lead to invalid results (Krull & MacKinnon, 1999, 2001; Raudenbush & Bryk, 2002).

An important form of multilevel mediation that is the focus of this paper is the two-level mediation model with one independent variable, one continuous mediator, and one continuous outcome variable, in which all of the variables are measured at level 1. Such a model is referred to as a 1–1–1 mediation model (Krull & MacKinnon, 1999, 2001). For ease of presentation, the 1–1–1 mediation model that we will consider in this paper has centred predictors and the intercepts and slopes are both random effects. Other types of multilevel mediation models exist as well (e.g., 2–1–1 for a model in which the independent variable is at level 2, and the mediator and outcome are at level 1), but are beyond the scope of the present paper (see Krull & MacKinnon, 2001; Preacher, Zyphur, & Zhang, 2010; Preacher, Zhang, & Zyphur, 2011).

3.1. Centring within contexts

Consider a two-level 1–1–1 mediation model example with a single mediator in which students (level 1) are nested within classrooms (level 2). In classroom j , X_{ij} indicates whether student i is randomly assigned to the treatment or control group, M_{ij} measures student i 's substance use refusal skill level, and Y_{ij} is student i 's self-report of the amount of substance use per day during the last 30 days (behaviour). Although the treatment indicator, mediator, and outcome variable are each measured at level 1, they may contain both between-cluster and within-cluster information. As MacKinnon (2008, p. 272) noted, '[m]any variables [at level 1] can be conceptualized at more than one level, making the clear interpretation of some multilevel models difficult'. To alleviate this problem, we use a centring strategy called centring within contexts 2 (CWC2; Enders & Tofighi, 2007; Kreft, de Leeuw, & Aiken, 1995; MacKinnon, 2008; Zhang, Zyphur, & Preacher, 2009). CWC2 uses CWC scores as a level-1 predictor and the cluster (observed or latent) means as a level-2 predictor of the random intercept in a multilevel analysis. Note that in practice one

can use either observed or latent cluster means in the computation of CWC2 scores (Lüdtke, Marsh, Robitzsch, Trautwein, Asparouhov, & Muthén, 2008; Preacher et al., 2010, 2011); each method has situations in which it is preferred as there is a trade-off between bias and efficiency (Lüdtke *et al.*, 2008). The analytic results presented in this paper hold for large sample sizes at levels 1 and 2 regardless of the centring method.² For simplicity, we use observed variable mean centring below and in the derivations in the Appendix. A more extensive presentation of the derivations is available in a separate online Appendix.

For a centred 1–1–1 mediation model we have the following level-1 equations:

$$Y_{ij} = d_{1j} + c_j(X_{ij} - \bar{X}_j) + \varepsilon_{1ij}, \quad (4)$$

$$M_{ij} = d_{2j} + a_j(X_{ij} - \bar{X}_j) + \varepsilon_{2ij}, \quad (5)$$

$$Y_{ij} = d_{3j} + c'_j(X_{ij} - \bar{X}_j) + b_j(M_{ij} - \bar{M}_j) + \varepsilon_{3ij}. \quad (6)$$

The level-2 (i.e., cluster-level) equations are as follows:

$$d_{1j} = d_1 + c_b \bar{X}_j + u_{d1j},$$

$$d_{2j} = d_2 + a_b \bar{X}_j + u_{d2j},$$

$$d_{3j} = d_3 + c'_b \bar{X}_j + b_b \bar{M}_j + u_{d3j},$$

$$c_j = c_w + u_{c_j},$$

$$a_j = a_w + u_{a_j},$$

$$c'_j = c'_w + u_{c'_j},$$

$$b_j = b_w + u_{b_j}$$

²Proofs of analytical results using latent variable centring are available upon request from the first author.

where the subscript w denotes a *within-cluster* effect and the subscript b denotes a *between-cluster* coefficient.

The level-1 and level-2 equations may be combined to form the mixed model equations for the centred 1–1–1 mediation model:

$$Y_{ij} = d_1 + c_b \bar{X}_j + c_w (X_{ij} - \bar{X}_j) + u_{d1j} + u_{c_j} (X_{ij} - \bar{X}_j) + \varepsilon_{1ij}, \quad (7)$$

$$M_{ij} = d_2 + a_b \bar{X}_j + a_w (X_{ij} - \bar{X}_j) + u_{d2j} + u_{a_j} (X_{ij} - \bar{X}_j) + \varepsilon_{2ij}, \quad (8)$$

$$Y_{ij} = d_3 + c'_b \bar{X}_j + b_b \bar{M}_j + c'_w (X_{ij} - \bar{X}_j) + b_w (M_{ij} - \bar{M}_j) + u_{d3j} + u_{c'_j} (X_{ij} - \bar{X}_j) + u_{b_j} (M_{ij} - \bar{M}_j) + \varepsilon_{3ij}, \quad (9)$$

where $\bar{X}_j$ and $\bar{M}_j$ are the observed means for cluster j . $(X_{ij} - \bar{X}_j)$ and $(M_{ij} - \bar{M}_j)$ are CWC scores for the treatment indicator and mediator variables, respectively. c_w is the within-cluster total effect, c_b is the between-cluster total effect, a_b is the between-cluster effect of X on M , a_w is the within-cluster effect of X on M , c'_b is the between-cluster direct effect, c'_w is the within-cluster direct effect, b_b is the between-cluster effect of M on Y controlling for X , and b_w is the within-cluster effect of M on Y controlling for X . The parameters d_1 , d_2 , and d_3 are the intercepts. The terms u_{d1j} , u_{d2j} , u_{d3j} , u_{c_j} , u_{a_j} , $u_{c'_j}$, and u_{b_j} are level-2 residuals for the intercepts and slopes. Figure 2 provides a graphical representation of the centred 1–1–1 mediation model. We generally follow the notation of Muthén and Muthén (1998–2010). A circle in the middle of an arrow represents a random slope, while a circle at the end of an arrow represents a random intercept. The relationships between the random intercepts are shown in the between-cluster section of the model.

3.2. Assumption: Orthogonality of residuals

We initially make the assumption that the centred 1–1–1 mediation model is properly specified – there is no omitted variable that causes the mediator and outcome variable to be related at either level 1 or level 2. This assumption is referred to as the *full orthogonality of residuals* assumption (McDonald, 1997). As we show below, the implication of this assumption is that the residuals for the level-1 and level-2 equations are (conditionally) uncorrelated. We limit our discussion to recursive models in both single-level and multilevel cases.

Full orthogonality is a standard assumption in *single-level* path and mediation models (Bollen, 1989; MacKinnon, 2008; McDonald, 1997). The full orthogonality assumption has two parts: first, the residual (error) terms in the model, here the mediation model in equations (2) and (3), are (mutually) uncorrelated; and second, the residual terms and exogenous variables are uncorrelated. The full orthogonality assumption implies that the postulated basic path or mediation model includes all variables that account for the covariances between the endogenous variables, here the covariance between M_i and Y_i . Further, it implies that, for a mediation model, there are no omitted variables in the model that can be considered a common cause of (a) a pair of exogenous and endogenous variables

in the model (e.g., X_j and M_j) or (b) a pair of endogenous variables in the model (e.g., M_j and Y_j).

For multilevel mediation models, the assumption of correct specification in terms of omitted variables is more complicated because residuals exist at both levels 1 and 2. An endogenous variable (e.g., M_{ij}) in a 1–1–1 mediation model in equations (8) and (9) has multiple residual terms compared to the single residual term in a single-level mediation model in equations (2) and (3). Extending the assumptions for single-level mediation models to the 1–1–1 mediation model, the full orthogonality assumption implies that the level-1 and level-2 residuals across equations (8) and (9) are uncorrelated. Finally, exogenous (independent) variables X_j and $(X_{ij} - \bar{X}_j)$ are independent of level-1 and level-2 residuals in equations (8) and (9). In addition, there can be no variables that are a common cause of each pair of X_j , $(X_{ij} - \bar{X}_j)$, M_j , $(M_{ij} - \bar{M}_j)$, and Y_j that are *not* included in the model.

3.3. Centred 1–1–1 mediation model with no omitted variables

Our first key result is that, given the full orthogonality assumption, the usual expression for single-level mediation, $c = a b + c'$, holds in a centred 1–1–1 mediation model for between- and within-cluster fixed effects separately. We sketch below the results of analytic work deriving the relationships between the population parameters in equations (7)–(9). The details of the derivation are presented in the Appendix.

In a centred 1–1–1 mediation model without an omitted variable, the following relationships between the coefficients in equation (7) and the coefficients in equations (8) and (9) hold:

$$c_b = c'_b + a_b b_b, \quad (10)$$

$$c_w = c'_w + a_w b_w.$$

Equation (10) shows that the CWC2 centring strategy partitions the total effect and the indirect effect into orthogonal *within-cluster* and *between-cluster* effects. The between-cluster indirect effect is defined as the between-cluster effect of the predictor on the outcome variable that is mediated by the mediator. The product of two coefficients, $a_b b_b$, measures the indirect effect at the aggregated cluster level. The within-cluster indirect effect is defined as the within-cluster effect of the independent variable on the outcome variable that is indirect through the mediator. The product of coefficients a_w and b_w quantifies the within-cluster indirect effect. In addition, these results indicate that the between-cluster total effect (c_b) is equal to the between-cluster indirect effect ($a_b b_b$) plus the between-cluster direct effect (c'_b). Similarly, the within-cluster total effect (c_w) is equal to the within-cluster indirect effect ($a_w b_w$) plus the within-cluster direct effect (c'_w).

Assuming that multivariate normality of the residuals holds, we arrive at the following relationships:

$$\sigma_{d_{1j}}^2 = \sigma_{d_{3j}}^2 + b_b^2 \sigma_{d_{2j}}^2,$$

$$\sigma_{c_j}^2 = \sigma_{c_j'}^2 + a_w^2 \sigma_{b_j}^2 + (b_w^2 + \sigma_{b_j}^2) \sigma_{a_j}^2 + 2a_w \sigma_{c_j'} \sigma_{b_j},$$

$$\sigma_{d_{1j}, c_j} = \sigma_{d_{3j}, c_j'} + a_w \sigma_{d_{3j}, b_j} + b_b b_w \sigma_{d_{2j}, a_j}. \quad (11)$$

3.4. Centred 1–1–1 multilevel mediation model with an omitted variable

We now explore what happens if the full orthogonality assumption does not hold. Randomized controlled trials rule out in expectation the possibility that unobserved variables might affect both the treatment assignment and the mediator or the treatment and the outcome. They do not rule out the possibility that unobserved variables might affect both the mediator and outcome variables. Consider a 1–1–1 mediation model in which a level-2 omitted variable W_j predicts between-cluster effects (i.e., W_j predicts the random intercepts) as well as serving as a moderator for the within-cluster effects (i.e., W_j predicts the random slopes). This model is shown in Figure 3. For simplicity, we investigate a case in which W_j is not expected to be correlated with treatment assignment X_{ij} , as would be true in a randomized trial. In our earlier example of a school-based drug abuse prevention programme, W_j might be a classroom characteristic such as the average grade point average (GPA) of a classroom. Suppose that classrooms with a higher GPA show a higher level of refusal skills and a lower level of substance abuse. In addition, we hypothesize that W_j moderates the level-1 relationships between the treatment, refusal skills and substance use.³ For example, the intervention might be more effective in increasing refusal skills and decreasing substance use in a classroom with a higher GPA compared to a classroom with a lower GPA.

Level-2 equations for the intercepts (between-cluster effects) are as follows:

$$d_{2j} = d_2 + a_b \bar{X}_j + h_{d_2} W_j + u_{d_{2j}},$$

$$d_{3j} = d_3 + c_b' \bar{X}_j + b_b \bar{M}_j + h_{d_3} W_j + u_{d_{3j}},$$

where h_{d_2} and h_{d_3} denote the fixed effects of W_j on the random intercepts associated with the mediator and outcome variable, respectively. Level-2 equations for random slopes (within-cluster effects) are as follows:

³In a non-randomized observational study, W_j could also affect the relationship between treatment assignment and the mediator and between treatment assignment and the outcome.

$$c_j = c_w + h_c W_j + u_{c_j},$$

$$a_j = a_w + h_a W_j + u_{a_j},$$

$$c'_j = c'_w + h_{c'} W_j + u_{c'_j},$$

$$b_j = b_w + h_b W_j + u_{b_j},$$

where h_c , h_a , $h_{c'}$, and h_b denote the fixed effects of W_j on the random slopes c_j , a_j , c'_j , and b_j , respectively. The mixed-effect equations are as follows:

$$Y_{ij} = d_1 + c_b \bar{X}_j + c_w (X_{ij} - \bar{X}_j) + h_{c'} w_j (X_{ij} - \bar{X}_j) + u_{d1j} + u_{c_j} (X_{ij} - \bar{X}_j) + \varepsilon_{1ij}, \quad (12)$$

$$M_{ij} = d_2 + a_b \bar{X}_j + h_{d2} W_j + a_w (X_{ij} - \bar{X}_j) + h_a W_j (X_{ij} - \bar{X}_j) + u_{d2j} + u_{a_j} (X_{ij} - \bar{X}_j) + \varepsilon_{2ij}, \quad (13)$$

$$\begin{aligned} Y_{ij} = & d_3 + c'_b \bar{X}_j + b_b \bar{M}_j + h_{d3} W_j + c'_w (X_{ij} - \bar{X}_j) + h_{c'} W_j (X_{ij} - \bar{X}_j) \\ & + b_w (M_{ij} - \bar{M}_j) + h_b W_j (M_{ij} - \bar{M}_j) + u_{d3j} + u_{c'_j} (X_{ij} - \bar{X}_j) \\ & + u_{b_j} (M_{ij} - \bar{M}_j) + \varepsilon_{3ij}. \end{aligned} \quad (14)$$

We are now in a position to analytically investigate the effect of omitting W_j on the between- and within-cluster effects (see Figure 4). The proofs of the following analytic results are shown in the Appendix. We denote the coefficients in the model with an omitted variable with an asterisk. First, we present the analytical results for the effect of the omitted variable on the between-cluster effects. To derive the between-cluster effects we use $W_j = l_0 + l_1 M_j + l_2 X_j + u_1$, where l_0 is an intercept, and l_1 and l_2 denote regression coefficients corresponding to the mediator and the independent variable, respectively; u_1 is a residual. The between-cluster fixed effects for the 1–1–1 mediation model with an omitted variable will be as follows:

$$a_b^* = a_b, \quad (15)$$

$$b_b^* = b_b + h_{d3} l_1, \quad (16)$$

$$c'_b{}^* = c'_b + h_{d3} l_2. \quad (17)$$

In addition, random intercept variances and covariance will be as follows:

$$\sigma_{d_{2j}}^2 = \sigma_{d_{2j}}^2 + h_{d_2}^2 \sigma_{W_j}^2,$$

$$\sigma_{d_{3j}}^2 = \sigma_{d_{3j}}^2 + h_{d_3}^2 \sigma_{u_1}^2,$$

$$\sigma_{d_{2j}^*, d_{3j}^*} = h_{d_2} h_{d_3} \sigma_{W_j}^2.$$

We next examine the effect of omitting a level-2 variable that serves as a moderator for the within-cluster effects. The within-cluster fixed effects will be as follows:

$$a_w^* = a_w + h_a \mu_{W_j}, \quad (18)$$

$$b_w^* = b_w + h_b \mu_{W_j}, \quad (19)$$

$$c_w'^* = c_w' + h_{c'} \mu_{W_j}. \quad (20)$$

The covariances and variances of the random slopes will be as follows:

$$\sigma_{a_j^*, b_j^*} = h_a h_b \sigma_{W_j}^2, \quad (21)$$

$$\sigma_{c_j'^*, b_j^*} = \sigma_{c_j', b_j} + h_{c'} h_b \sigma_{W_j}^2,$$

$$\sigma_{a_j^*, c_j'^*} = h_a h_{c'} \sigma_{W_j}^2,$$

$$\sigma_{a_j^*}^2 = \sigma_{a_j}^2 + h_a^2 \sigma_{W_j}^2,$$

$$\sigma_{b_j^*}^2 = \sigma_{b_j}^2 + h_b^2 \sigma_{W_j}^2,$$

$$\sigma_{c_j'^*}^2 = \sigma_{c_j'}^2 + h_{c'}^2 \sigma_{W_j}^2.$$

Finally, the effects of the omitted variable on the between-cluster and within-cluster total effects are as follows:

$$c_b^* = c_b' + a_b b_b, \quad (22)$$

$$c_w^* = c_w' + a_w b_w + (h_{c_j'} + h_a b_w + h_b a_w) \mu_{W_j} + h_a h_b \mu_{W_j}^2 + \sigma_{a_j^*, b_j^*}. \quad (23)$$

As can be seen in equations (22) and (23), under the assumptions considered here, the omitted variable does not affect the between-cluster total effect, whereas the within-cluster total effect will contain additional terms.

3.5. Interpretation and computation of indirect and total effects

The first key consequence of omitting a level-2 variable is that it affects the estimates of the between- and within-cluster fixed effects. For the between-cluster effects, the effect of the mediator on the outcome controlling for the effect of treatment assignment will be altered, whereas the effect of treatment assignment on the mediator as well as the total effect will be unchanged (see equations (15) and (22)). For the within-cluster indirect effect, the analytic results have an important substantive meaning given that the omitted variable is a *moderator*. In our substance use refusal skills example, the within-classroom indirect effect can now be interpreted as the *conditional* indirect effect for students who are in a classroom with a class GPA at the grand mean level (see Aiken & West, 1991). For students who are in a classroom with a different level of class GPA, the conditional indirect effect would be different. Given that GPA is omitted from the model, this result limits generalization of the within-cluster indirect effect. In addition, omitting the classroom GPA from the model leads the random slopes to become *correlated*. The larger the magnitude of the correlation between the random slopes, the greater the extent to which a classroom with a higher GPA compared to a classroom with a lower GPA will show the indirect effect. The prevention programme would lead to an increase in the student's level of refusal skills, which, in turn would lead to a decrease in drug use, which would be particularly strong in classrooms with a high GPA. The magnitude of the correlation between the random slopes depends on the heterogeneity of GPA between the classrooms in addition to the strength of relationships between the class GPA and the random slopes. Keeping other factors constant (e.g., the strength of relationships between the moderator and the random slopes), a higher variance in classroom GPA scores would result in a higher correlation between the random slopes.

A second, related, key consequence of a level-2 omitted variable is that the expected value of the product of random slopes ($a_j^* b_j^*$) and total effect (c_j^*) contains extra terms that account for the effect of the omitted variable. For the product of two random slopes, $a_j^* b_j^*$, the expected value contains an additional term (see equation (21)) that is a result of the omitted moderator. Equation (23) shows that the omitted variable will affect the estimate of the within-cluster total effect by the amount of $(h_{c_j'} + h_a b_w + h_b a_w) \mu_{W_j} + h_a h_b \mu_{W_j}^2 + \sigma_{a_j^*, b_j^*}$. The total within-cluster effect will include the total within-cluster effect of the independent

variable on the outcome variable, $a_w b_w + c'_w$, as well as the effect of the omitted variable. These results imply that a non-zero correlation between the random intercepts, random slopes, or both potentially indicates that there is an omitted variable at level 2.

3.6. Reconsidering the results of Kenny et al. (2003) and Bauer et al. (2006)

Kenny *et al.* (2003) and Bauer *et al.* (2006) proposed 1–1–1 mediation models that assumed correlations between a pair of random slopes associated with mediator and the outcome variables. Both Kenny *et al.* and Bauer *et al.* concluded that the expression $a + \sigma_{a_j, b_j} b_j$ quantified the indirect effect, with $c' + a + \sigma_{a_j, b_j} b_j$ equal to the total effect. However, as noted in the previous section, the correlation between the random slopes may be a result of an omitted variable that serves as a moderator for the within-cluster effects. For a centred 1–1–1 mediation model where the omitted variable is centred (for simplicity, $\mu_{W_j} = 0$), one implication of this analysis is that the within-cluster total effect in the expression

$c_w = a_w b_w + c'_w + \sigma_{a_j, b_j}$ is composed of the within-cluster direct and indirect effects plus a term that quantifies a covariance induced by the (centred) omitted variable. The within-cluster indirect effect of treatment assignment on the outcome conditional on the moderator being at the grand mean is $a_w b_w (E(a_j b_j | W_j = 0) = a_w b_w)$. Because there are interaction terms that involve the moderator, the interpretation of the indirect effect has become conditional. The additional term σ_{a_j, b_j} in the expression $E(a_j b_j) = a_w b_w + \sigma_{a_j, b_j}$ is a result of the interaction terms that involve the omitted variable. Similarly, the total within-cluster conditional on the grand mean of the omitted variable is $c_w - \sigma_{a_j, b_j} = a_w b_w + c'_w$, not

$c_w = a_w b_w + c'_w + \sigma_{a_j, b_j}$. The Kenny *et al.* and Bauer *et al.* models are more general than the basic 1–1–1 mediation model which assumes full orthogonality of residuals in that they permit misspecification at level 2. However, this generality comes at a cost of inducing the non-equivalence of the two methods of computing indirect effects and a more complicated interpretation of the within-cluster mediated effect as a conditional indirect effect.

3.7. Suggestions for practice

In model specification there is a tension between specification of the simpler, more interpretable model and specification of a more complex, but possibly less interpretable, more realistic model. The simpler model is often preferred as it represents the hypothesized theoretical model in the area. Such a representation is particularly important in the case of mediational analysis where the researchers' goal is often to test a hypothesized causal mechanism; such tests currently require strong assumptions (Holland, 1988; Imai, Keele, & Yamamoto, 2010; West, 2011). At the same time, models that do not adequately represent the data can easily lead to biased estimates of key effects (e.g., Cole, Ciesela, & Steiger, 2007), here the within-cluster indirect effect. Below, we offer a procedure for practising researchers that attempts to balance these two competing tensions.

We recommend that researchers estimate two 1–1–1 CWC2 mediation models (a) without and (b) with random effects associated with M and Y correlated. If a significant covariance between a_j and b_j is not found, the full orthogonality of residuals assumption is plausible and the basic multilevel mediation model should be retained given its simpler interpretation. If a significant covariance between a_j and b_j is found, the full orthogonality assumption is

implausible, and the alternative multilevel mediation model proposed by Kenny *et al.* (2003) and Bauer *et al.* (2006) should be considered. The next step is to explore potential candidates for the level-2 omitted variables and include them in the model. If one can find the moderators that account for the covariance between a_j and b_j , then estimating and interpreting the indirect effect is straightforward, as described in Bauer *et al.* (2006). In this case, we still recommend using the formula $c_w = c'_w + a_w b_w$ to calculate the total effect at various levels of moderator rather than estimating it directly because of the presence of the quadratic term W_j^2 in equation (A6).

If after including potential moderators the term $\sigma_{a_j b_j}$ is still significant, then we recommend that a sensitivity analysis be conducted. In this sensitivity analysis, the covariance between the random slopes can be reparametrized as a latent variable with the mean zero and standard deviation one. For identification purposes, the effects of latent variable on the random slopes are assumed to be equal. The latent variable serves as a proxy for the omitted variable at level 2. To probe the moderated indirect effects, for example, one can choose values of -1 , 0 , and $+1$ for the latent variable to estimate conditional within-cluster indirect effects at the levels of the latent variable that correspond to mean and mean plus or minus one standard deviation. When the latent variable equals 0 , the terms $a_w b_w$ and $c_w = c'_w + a_w b_w$ represent the within-cluster indirect effect and total effect for a cluster whose value equals the grand mean of the latent moderator. For the reasons stated previously, we recommend computing the value of the total effect c_w indirectly from the formula $c_w = c'_w + a_w b_w$ instead of estimating it directly from equation (12).

4. Simulation

A small-scale simulation study was conducted to empirically investigate the effects of a level-2 omitted variable on the fixed effects in a 1–1–1 mediation model. Data generation was based on the population model for the 1–1–1 mediation model shown in equations (13) and (14) with a level-2 predictor W_j for both intercepts and slopes. The population values for the model were as follows: $c'_b = 0.15$, $a_b = 0.59$, $b_b = 0.15$, $c_b = 0.2385$, $c'_w = 0.15$, $a_w = 0.39$, $b_w = 0.39$, $c_w = 0.3021$, $h_a = 0.4$, $h_b = 0.5$, $h_c = 0.1$, $h_{d2} = 0.3$, $h_{d3} = 0.4$, $\mu_{W_j} = 1$, $\sigma_{W_j}^2 = 1$, $\mu_{X_j} = 2$ and $\sigma_{X_j}^2 = 1$. The population values for the between- and within-cluster total effects were inferred parameters based on equation (10). For simplicity, level-1 and level-2 variances were all set to 1; level-2 covariances as well as the intercepts were fixed at 0. The simulation design included five conditions varying the number of clusters (50, 100, 200, 500, and 1000); the cluster size was always 10. For each condition, we generated 1000 data sets (replications). Simulated data were analysed using the model in which W_j was omitted. The multilevel analysis was performed using the package *lme4* (Bates, Mächler, & Bolker, 2011) in R (R Development Core Team, 2011). Table 1 presents the results of the simulation. Consistent with the analytical results shown in equations (15)–(20), (22) and (23), the omitted variable affected the estimates of within- and between-cluster effects except for the parameters a_b and c_b . For some parameters the magnitude of the effect was large, with the estimates being more than twice the size of the respective population values. The effect of

the omitted variable on the parameter estimates depends on the mean and variance of the omitted variable as well as on the covariance between the omitted variable and mediator and the outcome variable.

5. Summary and conclusion

Many data collection efforts occur in group settings (e.g., classroom or community) or involve repeated measurements over time, giving rise to a need for multilevel mediation models that take clustering into account. In this paper, we initially derived the algebraic relationship between the population parameters in a CWC2 1–1–1 mediation model with no omitted variables. When it was assumed that the basic level-1 and level-2 equations were correctly specified in terms of no omitted variables, this implied that the level-1 residuals and level-2 residuals would *not* be correlated across equations (8) and (9) (i.e., the full orthogonality assumption). The full orthogonality assumption led to the important consequence that $c_b = a_b b_b + c'_b$ and $c_w = a_w b_w + c'_w$.

We also examined the effect of a level-2 omitted variable that affects both the mediator and the outcome variables. We analytically showed that the omitted variable can affect both the between- and within-cluster fixed effect estimates. In addition, an omitted variable that serves as moderator for the within-cluster effects causes the random slopes associated with the mediator and outcome variable to be correlated. Note that this is a spurious correlation that is an artefact of the omitted moderator. For ease of presentation, we assumed in deriving the analytical results for the 1–1–1 mediation model that the omitted variable was *not* correlated with the independent variable, as would be the case in a randomized experiment. This assumption can be relaxed. In such cases the amount of bias in the between-cluster parameter estimates will additionally depend on the between-cluster correlation between the omitted variable and the independent, mediator, and outcome variables. This increased complexity and possibly realism comes at a cost of making the within-cluster indirect effect conditional so that it applies to clusters whose mean values on the omitted variable are close to the grand mean. It also substantially increases the challenges of making causal inferences about any mediation effects (West, 2011). Another finding of our derivation was that, while the CWC2 centring strategy makes the between-cluster and within-cluster effects theoretically orthogonal, it did *not* account for the correlation between level-2 residuals (random coefficients).

The present results imply that a non-zero correlation between random intercepts can be a potential indicator of an omitted variable at level 2 that is a common predictor of the random intercepts, and a non-zero correlation between random slopes can be an indicator of an omitted variable at level-2 that serves as a moderator for the within-cluster effects. The formulation by Kenny *et al.* (2003) and Bauer *et al.* (2006) of the indirect effect and total effect potentially includes the effect of an omitted variable in addition to the indirect and total effect of the independent variable on the outcome variable. The corrected formulas for the conditional (at the grand mean of the omitted variable) within-cluster indirect and total effect are $a_w b_w$ and $c_w - \sigma_{a_j, b_j} = c'_w + a_w b_w$, respectively.

The results of the present analytic work are limited to two-level 1–1–1 mediation models which are common multilevel mediation models. Our analytic results can be extended to structural equation modeling approaches to the two-level 1–1–1 mediation model given the general equivalence of the structural equation and multilevel model approaches with observed variables (Curran, 2003). A variety of other multilevel mediation models in which data collection for at least one of the variables takes place at level 2 (e.g., 2–1–1, 2–2–1) were not considered. Note that in these other two-level mediation models, the within-cluster indirect effect cannot be properly estimated because at least one of the variables in the model (X , M , or Y) will *not* vary within clusters. In these models, only the between-cluster indirect effects can be estimated (Zhang *et al.*, 2009). In addition, the results of the present work are limited to multilevel mediation models in which variables are measured at two levels. Multilevel mediation models with three levels can also be considered. These models remain topics for future research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The research was supported by Award Number F31DA027336 to Davood Tofighi and PHS DA09757 to David P. MacKinnon from the National Institute on Drug Abuse. Stephen G. West was supported by a Forschungspreis from the Alexander von Humboldt foundation.

References

- Aiken, LS.; West, SG. Multiple regression: Testing and interpreting interactions. Thousand Oaks, CA: Sage; 1991.
- Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*. 1986; 51:1173–1182. [PubMed: 3806354]
- Bates DM, Mächler M, Bolker BM. lme4: Linear Mixed-Effects Models Using S4 Classes (Version 0.999375–39) [Software]. 2011 Retrieved from <http://CRAN.R-project.org/package=lme4>.
- Bauer DJ, Preacher KJ, Gil KM. Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods*. 2006; 11:142–163. [PubMed: 16784335]
- Bollen, K. Structural equations with latent variables. New York: Wiley; 1989.
- Botvin GJ, Eng A. A comprehensive school-based smoking prevention program. *Journal of School Health*. 1980; 50:209–213. [PubMed: 6900141]
- Cole DA, Ciesla JA, Steiger JH. The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*. 2007:381–398. [PubMed: 18179350]
- Curran PJ. Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*. 2003; 38:529–569. [PubMed: 26777445]
- Ellickson PL, McCaffrey DF, Ghosh-Dastidar B, Longshore DL. New inroads in preventing adolescent drug use: Results from a large-scale trial of project ALERT in middle schools. *American Journal of Public Health*. 2003; 93:1830–1836. [PubMed: 14600049]
- Enders CK, Tofighi D. Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*. 2007; 12:121–138. [PubMed: 17563168]

- Hawkins JD, Brown EC, Oesterle S, Arthur MW, Abbott RD, Catalano RF. Early effects of communities that care on targeted risks and initiation of delinquent behavior and substance use. *Journal of Adolescent Health*. 2008; 43:15–22. [PubMed: 18565433]
- Holland, PW. Causal inference, path analysis, and recursive structural equation models. In: Clogg, CC., editor. *Sociological methodology*. Washington, DC: American Psychological Association; 1988. p. 449–493.
- Imai K, Keele L, Yamamoto T. Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*. 2010; 25:51–71.
- Kenny DA, Korchmaros JD, Bolger N. Lower level mediation in multilevel models. *Psychological Methods*. 2003; 8:115–128. [PubMed: 12924810]
- Kreft IGG, de Leeuw J, Aiken LS. The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*. 1995; 30:1–21. [PubMed: 26828342]
- Krull JL, MacKinnon DP. Multilevel mediation modeling in group-based intervention studies. *Evaluation Review*. 1999; 23:418–444. [PubMed: 10558394]
- Krull JL, MacKinnon DP. Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research*. 2001; 36:249–277. [PubMed: 26822111]
- Lüdtke O, Marsh HW, Robitzsch A, Trautwein U. A 2×2 taxonomy of multilevel latent contextual models: Accuracy–bias trade-offs in full and partial error correction models. *Psychological Methods*. 2011:444–467. [PubMed: 21787083]
- Lüdtke O, Marsh HW, Robitzsch A, Trautwein U, Asparouhov T, Muthén B. The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*. 2008; 13:203–229. [PubMed: 18778152]
- Lynagh M, Schofield MJ, Sanson-Fisher RW. School health promotion programs over the past decade: A review of the smoking, alcohol and solar protection literature. *Health Promotion International*. 1997; 12:43–60.
- MacKinnon, DP. *Introduction to statistical mediation analysis*. New York: Erlbaum; 2008.
- MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annual Review of Psychology*. 2007; 58:593–614.
- MacKinnon DP, Warsi G, Dwyer JH. A simulation study of mediated effect measures. *Multivariate Behavioral Research*. 1995; 30:41–62. [PubMed: 20157641]
- McDonald RP. Haldane's lungs: A case study in path analysis. *Multivariate Behavioral Research*. 1997; 32:1–38. [PubMed: 26751104]
- Muthén, LK.; Muthén, BO. *Mplus user's guide*. 6th. Los Angeles: Muthén & Muthén; (1998–2010).
- Preacher KJ, Zhang Z, Zyphur MJ. Alternative methods for assessing mediation in multilevel data: The advantages of multilevel SEM. *Structural Equation Modeling*. 2011; 18:161–182.
- Preacher KJ, Zyphur MJ, Zhang Z. A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*. 2010; 15:209–233. [PubMed: 20822249]
- R Development Core Team. Vienna, Austria: R Foundation for Statistical Computing; 2011. R: A language and environment for statistical computing (Version 2.13.0). Retrieved from <http://www.R-project.org/>
- Raudenbush, SW.; Bryk, AS. *Hierarchical linear models : Applications and data analysis methods*. 2nd. Thousand Oaks, CA: Sage; 2002.
- West SG. Editorial: Introduction to the special section on causal inference in cross sectional and longitudinal mediational models. *Multivariate Behavioral Research*. 2011; 46:812–815. [PubMed: 26736046]
- Zhang Z, Zyphur MJ, Preacher KJ. Testing multilevel mediation using hierarchical linear models. *Organizational Research Methods*. 2009; 12:695–719.

a) Total Effect

b) Indirect or Mediated Effect

Figure 1.
A basic single-mediator model.

Figure 2.

Within- and between-cluster effects for a centred 1–1–1 mediation model with no level-2 predictor. The model includes both random intercepts and slopes. Level-1 residuals are not shown. Residuals associated with the mediator and outcome variables are not correlated. u_{b_j} and $u_{c'_j}$ are correlated but not shown.

Figure 3.

Within- and between-cluster effects for a centred 1-1-1 mediation model with W_j as a level-2 predictor of the random intercepts and slopes. Level-1 residuals are not shown.

Residuals associated with the mediator and outcome variables are not correlated. u_{b_j} and $u_{c'_j}$ are correlated but not shown.

Figure 4.

Within- and between-cluster effects for a centred 1-1-1 mediation model with an omitted variable at level-2. Dashed curved lines represent the 'spurious' correlations caused by an omitted variable W_j that is a common predictor of the random intercepts and slopes. Level-1 residuals are not shown. u_{b_j} and $u_{c'_j}$ are correlated but not shown.

Table 1
Percentage bias for the centred 1–1–1 mediation model with a level-2 omitted variable (cluster size = 10)

Parameter	Population	Number of clusters											
		50			100			200			500		
		Bias	RMSE		Bias	RMSE		Bias	RMSE		Bias	RMSE	
c_b	0.15	-26.8%	0.198		-36.6%	0.131		-37.9%	0.093		-34.4%	0.055	
a_b	0.59	-0.5%	0.163		0.5%	0.107		-0.7%	0.079		-0.4%	0.050	
b_b	0.15	55.7%	0.156		61.8%	0.114		62.6%	0.077		60.3%	0.047	
c_b	0.2385	2.4%	0.188		-0.2%	0.123		-1.4%	0.088		0.6%	0.055	
c_w	0.15	74.0%	0.161		75.3%	0.117		76.0%	0.081		76.4%	0.052	
a_w	0.39	102.8%	0.156		103.1%	0.113		102.5%	0.079		103.1%	0.049	
b_w	0.39	131.7%	0.174		125.9%	0.118		129.8%	0.081		129.7%	0.054	
c_w	0.3021	287.0%	0.320		279.6%	0.226		284.7%	0.159		281.4%	0.098	

Note. The numbers in shaded rows correspond to results for the inferred parameters c_b and c_w . $RMSE$ = root mean square error.