

# The Effects of Misspecifying the Random Part of Multilevel Models

David M. LaHuis<sup>a</sup>, Daniel R. Jenkins<sup>a</sup>, Michael J. Hartman<sup>b</sup>, Shotaro Hakoyama<sup>b</sup>,  
Patrick C. Clark<sup>b</sup>

[a] *Department of Psychology, Wright State University, Dayton, OH, USA.* [b] *Infor, Dallas, TX, USA.*

---

Methodology, 2020, Vol. 16(3), 224–240, <https://doi.org/10.5964/meth.2799>

**Received:** 2019-03-26 • **Accepted:** 2019-10-16 • **Published (VoR):** 2020-09-30

**Corresponding Author:** David M. LaHuis, Department of Psychology, Wright State University, Dayton, OH 45435-0001, USA. Tel. +1 937 775 3818, E-mail: david.lahuis@wright.edu

---

## Abstract

This paper examined the amount bias in standard errors for fixed effects when the random part of a multilevel model is misspecified. Study 1 examined the effects of misspecification for a model with one Level 1 predictor. Results indicated that misspecifying random slope variance as fixed had a moderate effect size on the standard errors of the fixed effects and had a greater effect than misspecifying fixed slopes as random. In Study 2, a second Level 1 predictor was added and allowed for the examination of the effects of misspecifying the slope variance of one predictor on the standard errors for the fixed effects of the other predictor. Results indicated that only the standard errors of coefficient relevant to that predictor were impacted and that the effect size for the bias could be considered moderate to large. These results suggest that researchers can use a piecemeal approach to testing multilevel models with random effects.

## Keywords

multilevel, random effects, fixed effects, Monte Carlo

The use of multilevel modeling has become common in social science research because it allows researchers to analyze data that reside at different levels. Multiple levels may include employees nested within teams, students nested within classrooms, or measurements nested within individuals. A strength of multilevel modeling is the flexibility that it allows researchers in testing how variables at higher levels relate to variables at lower levels. A cost of the flexibility is that multilevel models can be complicated to specify. As such, studies have investigated the effects of misspecifying multilevel models. For example, studies have investigated the effects of ignoring (a) the heterogeneity of Level 2 variances (Korendijk, Maas, Moerbeek, & Van der Heijden, 2008), (b)



inequalities of within-group effects (Korendijk, Hox, Moerbeek, & Maas, 2011), (c) a level of nesting (Moerbeek, 2004), (d) random cross classifications (Gilbert, Petscher, Compton, & Schatschneider, 2016; Luo & Kwok, 2009; Meyers & Beretvas, 2006; Shi, Leite, & Algina, 2010), (e) multiple membership structures (Chung & Beretvas, 2012), and (f) partial clustering (Baldwin, Bauer, Stice, & Rohde, 2011). The effects of these misspecifications primarily affect the standard errors of the regression coefficients and variance components.

In multilevel modeling, specifying the random part of the model is often the most difficult because there are rarely strong theories about which effects should vary. Proper specification of the random part of the model is important because it influences the standard errors of the fixed effects (Snijders & Bosker, 2011). For example, failing to model random slopes will decrease the standard errors and modeling nonrandom slopes may lead to an overly complicated model resulting in estimation problems.

Despite its importance, there has been little research concerning the effects of misspecifying the random part of the model. One exception is Berkhof and Kampen (2004) who derived expressions for the changes in the standard errors for fixed effects as a result of omitting a random effect in a multilevel model. They based the expressions on moment estimators that yield closed form expressions. Although they found that their expressions were useful in determining the general effect of omitting a random effect, the expressions were not useful in describing the magnitude of the effect. In addition, they focused on a single predictor and did not examine the effects of including an extra random effect. Thus, there are still some questions about the effects of misspecifying the random part of the model.

In the present studies, we examined the consequences of misspecifying the random part of the model on the standard errors of the fixed effects. We were interested in the effects of specifying fixed slopes as varying and varying slopes as fixed. Because closed-form expressions are not possible when there are unequal group sizes or random slopes, we used Monte Carlo simulations to assess the effects of model misspecification. In our simulations, we assessed bias by comparing standard errors for the fixed effects from a model where the slope varied to a model where the slope was fixed. Our goal was to determine which type of misspecification produced the most detrimental effects on the standard errors of the fixed effects.

## Multilevel Modeling

A succinct way of expressing a basic two-level model where there are  $P$  Level 1 predictors ( $X$ 's) and  $Q$  Level 2 predictors ( $Z$ 's) is:

$$Y_{ij} = \gamma_{00} + \sum_{h=1}^P \gamma_{h0} X_{hij} + \sum_{k=1}^Q \gamma_{0k} Z_{kj} + \sum_{k=1}^Q \sum_{h=1}^P \gamma_{hk} Z_{kj} X_{hij} + u_{0j} + \sum_{h=1}^P u_{hj} X_{hij} + r_{ij} \quad (1)$$

In Equation 1, the Level 1 error is labeled  $r_{ij}$  and its variance ( $\sigma^2$ ) represents within group variance not explained by the model. The subscripts  $i$  and  $j$  refer to the Level 1 and 2 units, respectively. There are  $P + 1$  random parameters because every Level 1 predictor has a random slope and there is one random intercept. In practice, only a subset of Level 1 predictors may be specified as having random slopes. The first four terms contain the fixed part of the model, and the last three terms contain the random part of the model. The variance of  $u_{0j}$  is represented as  $\tau_{00}$  and represents intercept variance not explained by the  $Z$ s. Similarly, the variances of  $u_{hj}$  are labeled  $\tau_{hh}$  and represent the unexplained variances in the Level 1 slopes. It is common to allow the intercept and slopes to covary.

Researchers are often advised to test for the significance of the variance components in order to determine what effects should vary. One potential problem with using significance tests for the slope variance components is that the power for these tests is low (LaHuis & Ferguson, 2009; Raudenbush & Liu, 2000). Thus, these tests have greater likelihoods of committing Type II errors. The result of such an error is that the random part of the multilevel model will be misspecified in that the slope is fixed when it should vary. As a result, the standard errors for the fixed effects would be too small. This would lead to a greater risk of Type I errors for the fixed effects.

### Random Effects and Standard Errors

To understand how the misspecification of the random part of the model affects the standard errors for the fixed effects, it is helpful to consider the unrealistic case where the variance components are known, every Level 1 predictor has a random slope, and each group has sufficient size to calculate ordinary least squares (OLS) estimates. Following Raudenbush and Bryk (2002, Chapter 3), the Level-1 model for group  $j$  in an example with  $P$  Level 1 predictors can be expressed as

$$Y_j = X_j\beta_j + r_j \quad (2)$$

where  $Y_j$  is a vector of  $n_j$  outcome scores.  $X_j$  is a  $n_j$  by  $(P + 1)$  matrix where the first column is a series of ones to represent the intercept and the other columns are the Level 1 predictors.  $\beta_j$  is a vector of unknown parameters and  $r_j$  are the random errors with a mean of zero and a variance of  $\sigma^2$ . The ordinary least squares (OLS) estimator for the unknown parameters is

$$\hat{\beta}_j = (X_j^T X_j)^{-1} X_j^T Y_j \quad (3)$$

The error variance matrix for  $\hat{\beta}_j$  in estimating  $\beta_j$  is denoted as  $V_j$  and is given by

$$V_j = \sigma^2 (X_j^T X_j)^{-1} \quad (4)$$

Finally, pre-multiplying both sides of Equation 2 by  $(X_j^T X_j)^{-1} X_j^T$  produces a simpler model for  $\hat{\beta}_j$ :

$$\hat{\beta}_j = \beta_j + e_j, \quad e_j \sim N(0, V_j) \quad (5)$$

At Level 2, the model for  $\beta_j$  is

$$\beta_j = Z_j \gamma + u_j, \quad u_j \sim N(0, T) \quad (6)$$

In Equation 6,  $\gamma$  is a  $F$  by 1 vector of fixed effects, the matrix  $Z_j$  is a  $(P + 1)$  by  $F$  matrix that is stacked in block diagonal fashion. For example, in a model with one Level 2 variable  $Z_j$  predicting one Level 1 intercept and one Level 1 slope, Equation 6 would be:

$$\begin{pmatrix} \beta_{0j} \\ \beta_{1j} \end{pmatrix} = \begin{pmatrix} 1 & Z_j & 0 & 0 \\ 0 & 0 & 1 & Z_j \end{pmatrix} \begin{pmatrix} \gamma_{00} \\ \gamma_{01} \\ \gamma_{10} \\ \gamma_{11} \end{pmatrix} + \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \quad (7)$$

Both  $u_{0j}$  and  $u_{1j}$  are the random effects representing the intercept and slope, respectively. They are assumed to be normally distributed with a mean vector of zeros and a variance covariance matrix  $\mathbf{T}$ .

Substituting Equation 6 into Equation 5 creates a single combined model:

$$\hat{\beta}_j = Z_j \gamma + u_j + e_j \quad (8)$$

The dispersion of  $\hat{\beta}_j$  conditional on  $Z_j$  is a function of the error variance matrix at Level 1 and the random effects matrix at Level 2:

$$\text{var}(\hat{\beta}_j) = \Delta_j = T + V_j = \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{01} & \tau_{11} \end{pmatrix} + \begin{pmatrix} V_{00j} & V_{01j} \\ V_{01j} & V_{11j} \end{pmatrix} = \begin{pmatrix} \tau_{00} + V_{00j} & \tau_{01} + V_{01j} \\ \tau_{01} + V_{01j} & \tau_{11} + V_{11j} \end{pmatrix} \quad (9)$$

Finally, the dispersion matrix for the fixed effects ( $\gamma$ 's) based on generalized least squares is given by:

$$\text{var}(\hat{\gamma}_j) = V_\gamma = \left( \sum_{j=1}^J (Z_j^T \Delta_j^{-1} Z_j) \right)^{-1} \quad (10)$$

The square roots of the diagonals of  $V_\gamma$  are the standard errors. To trace the effects of  $\Delta_j$  on the standard errors, it is helpful to consider a very simple model with two Level 1 predictors and no Level 2 predictors. The underlying concepts are identical for models

with more fixed and random effects but are more difficult to demonstrate. Thus, in our simple model,  $Z$  is simply a  $3 \times 3$  identity matrix and Equation 10 becomes:

$$\begin{aligned} \text{var}(\hat{Y}_j) = V_Y &= \left( \sum_{j=1}^J (Z_j^T \Delta_j^{-1} Z_j^T) \right)^{-1} \\ &= \left( \sum_{j=1}^J \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tau_{00} + V_{00j} & \tau_{01} + V_{01j} & \tau_{02} + V_{02j} \\ \tau_{01} + V_{01j} & \tau_{11} + V_{11j} & \tau_{12} + V_{12j} \\ \tau_{02} + V_{02j} & \tau_{12} + V_{12j} & \tau_{22} + V_{22j} \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right)^{-1} \end{aligned} \quad (11)$$

standard errors for  $\gamma_{00}$ ,  $\gamma_{10}$ , and  $\gamma_{20}$  are simply the square roots of the diagonals of  $V_j$ . Importantly, the slope variance for one Level 1 predictor does not impact the standard error for the other Level 1 predictor. It follows that fixing the slope of a given predictor when it should vary should only affect the standard errors associated with that predictor.

In practice, the variance components are not known and need to be estimated with the fixed effects using full or restricted maximum likelihood estimation. Based on the above discussion, we would expect the bias in standard errors would be limited to those fixed effects that are associated with the random effect that is misspecified and that the amount of this bias should decrease as the amount of slope variance decreases. In addition, the amount of bias in standard errors should increase as group size increases. This is because, all things being equal, the elements of the error matrix  $V_j$  will decrease as group size increases. Thus, the Level 2 variances will have a greater relative impact on the standard errors of the regression coefficients.

## Study 1

The major focus of Study 1 was to examine the effects of omitting a slope from the random part of the model. We also explored the effects of freeing a slope when it should have been fixed. We considered a simple model with only one Level 1 predictor and one Level 2 predictor. As mentioned above, specification of the random part of the model impacts the standard errors for the fixed effects. We assessed the effect of misspecification on the standard errors of three fixed effects: the cross-level effect ( $\gamma_{01}$ ), the slope mean for  $X$  ( $\gamma_{10}$ ), and the cross-level interaction ( $\gamma_{11}$ ). Misspecification occurred by allowing the slope to vary when it should have been fixed and vice-versa. For each fixed effect, we were interested in the overall degree of bias. In addition, we assessed the impact of the number of groups, average group size, the mean slope, the magnitude of the cross-level interaction, and the amount of slope variance on the amount of bias.

## Method

We used a 3 (number of groups)  $\times$  3 (average group size)  $\times$  2 (cross-level interaction effect size)  $\times$  4 (slope mean)  $\times$  3 (slope – intercept correlation)  $\times$  5 (slope variance effect size) design. The values chosen for the conditions were based somewhat on the simulations conducted by Mathieu, Aguinis, Culpepper, and Chen (2012). The intraclass correlation coefficient (ICC) for  $Y$  was set to .20, and the cross-level effect was set at 0.40. The number of groups was 50, 100, or 200. The average group sizes were 5, 10, and 20, and were chosen to be consistent with previous simulations (LaHuis & Ferguson, 2009; LaHuis, Hartman, Hakoyama, & Clark, 2014; Mathieu et al., 2012). Group sizes were drawn randomly from a uniform distribution. For the average group size of 5 condition, values ranged from 1 to 9. The range of group sizes for average group size of 10 condition was 6 to 14. In the average group size of 20 condition, the range was 16 to 24. We also included a cross-level interaction condition where the effect size for the cross-level interaction was either 0.00 or 0.20. The latter value was based on Mathieu et al (2012). The slope means were specified as 0.00, 0.10, 0.30, or 0.50. These represent nil, small, medium, and large effects, respectively. Slope variance values were 0.00, 0.025, 0.05, 0.10, and 0.15 which produced 95% predictive intervals (Hox, Moerbeek, & van de Schoot, 2018) of the slope mean plus or minus 0, .31, .44, .62, .76, respectively. That is, for the slope mean condition of 0.10 and the slope variance condition of 0.05, we would expect 95% of the slopes to be in the range of -.34 to .54. The slope variance values were consistent with Raudenbush and Liu (2000) and represent a range of effect sizes for slope variances. Finally, the correlation between the slope and intercept was set to .00, -.30, or -.70. For each possible combination, we simulated 1000 samples resulting in 936,000 samples.

## Data Generation

We created the Level 2 variable  $Z$  by randomly sampling values from a standard normal distribution. We also generated values for the Level 2 random variables by randomly sampling values from a normal distribution with a mean of 0 and a standard deviation equal to the square root of the desired variance. For  $u_{0j}$ , the standard deviation was set to the value to produce an ICC for  $Y$  of .20. The value changed based on the amount of explained variance at each level. For  $u_{1j}$ , the standard deviation was set to the value to produce the desired slope variance. At Level 1, we randomly sampled values from a standard normal distribution to produce the Level 1 predictor  $X$ . In addition, the Level 1 error term,  $r_{ij}$ , was created by randomly drawing values from a normal distribution with a mean of 0 and a standard deviation set to produce the desired ICC for  $Y$ .

To generate the  $Y$ , we first created intercepts for each group by multiplying  $Z$  by 0.40 (the cross-level effect) and added  $u_{0j}$ . Similarly, we created the group slopes ( $\beta_{1j}$ ) by adding the desired mean slope (i.e., 0.00, 0.10, 0.30, or 0.50) and the product of  $Z$  with

the cross-level interaction effect to  $u_{1j}$ . Finally, we computed values for  $Y$  using the group intercept and slope values and the randomly generated values for  $X$  and  $r_{ij}$ .

## Analyses

For each sample, we estimated two multilevel models using restricted maximum-likelihood estimation: one where the slope was fixed and one where the slope varied. For each model, we recorded the estimates for  $\gamma_{01}$ ,  $\gamma_{10}$ , and  $\gamma_{11}$ , as well as their standard errors. We calculated the bias in standard errors by subtracting the standard error from the data generating model from that of the misspecified model. Thus, a positive value for bias reflects the standard errors for the misspecified model was higher than the data generating model. Cohen's  $d$  statistics were also calculated to aid in interpreting the size of the bias. We used the guidelines (Cohen, 1988) of .20, .50, and .80 as indicative of small, medium, and large effects, respectively. We conducted ANOVA's to determine the impact of the various conditions on the bias in standard errors. Because of the large number of samples, we focused on interpreting effect sizes (partial  $\eta^2$ ). All analyses were conducted using the LME4 (Bates, Maechler, Bolker, & Walker, 2015) package in the statistical program R (R Core Team, 2018).

## Results

Prior to evaluating the results, we inspected how well the population parameters were uncovered by our analyses. We calculated root mean square error of approximation values (RMSEA) for each parameter. RMSEA values ranged from .02 to .03. Thus, all values were in the acceptable range and consistent with prior simulations.

### Bias in Standard Errors

Table 1 presents the average bias statistics for each condition. These statistics were calculated by collapsing across the other conditions. Misspecification of the random part of the model did not appear to impact  $\gamma_{01}$ . The Cohen's  $d$  statistics ranged from -0.02 to 0.02. In addition, the bias values were small: They were all near .000.

**Table 1**

*Bias Statistics for the Standard Error Differences in the Parameter Estimates Between a Model With a Fixed Slope and a Model With a Slope Allowed to Vary*

Source	$\gamma_{01}$			$\gamma_{10}$			$\gamma_{11}$		
	SE	SE Bias	$d$	SE	SE Bias	$d$	SE	SE Bias	$d$
NJ = 5	.090	.000	0.01	.061	-.009	-0.52	.062	-.009	-0.47
NJ = 10	.079	.000	0.01	.046	-.010	-0.77	.046	-.010	-0.72
NJ = 20	.075	.000	0.01	.037	-.012	-1.14	.037	-.012	-1.08

Source	$\gamma_{01}$			$\gamma_{10}$			$\gamma_{11}$		
	SE	SE Bias	<i>d</i>	SE	SE Bias	<i>d</i>	SE	SE Bias	<i>d</i>
NG = 50	.110	.000	0.02	.066	-.014	-0.84	.067	-.014	-0.76
NG = 100	.078	.000	0.02	.046	-.010	-0.87	.046	-.010	-0.82
NG = 200	.055	.000	0.02	.032	-.007	-0.88	.033	-.007	-0.85
$\gamma_{10} = 0.00$	.081	.000	0.01	.048	-.010	-0.59	.049	-.011	-0.56
$\gamma_{10} = 0.10$	.081	.000	0.01	.048	-.010	-0.59	.049	-.011	-0.56
$\gamma_{10} = 0.30$	.081	.000	0.01	.048	-.010	-0.59	.049	-.010	-0.56
$\gamma_{10} = 0.50$	.081	.000	0.01	.048	-.010	-0.59	.049	-.010	-0.56
$\tau_{11} = 0.00$	.044	.000	-0.02	.035	.001	0.09	.036	.001	0.09
$\tau_{11} = 0.025$	.084	.000	0.00	.041	-.005	-0.28	.042	-.005	-0.26
$\tau_{11} = 0.05$	.084	.000	0.01	.045	-.008	-0.47	.046	-.008	-0.44
$\tau_{11} = 0.10$	.084	.000	0.01	.052	-.014	-0.78	.052	-.014	-0.74
$\tau_{11} = 0.15$	.084	.001	0.02	.058	-.019	-1.03	.058	-.019	-0.96
$\tau_{01} = 0.00$	.076	.000	0.01	.047	-.009	-0.53	.047	-.009	-0.50
$\tau_{01} = -0.30$	.084	.000	0.01	.049	-.011	-0.64	.050	-.011	-0.60
$\tau_{01} = -0.70$	.084	.000	0.01	.049	-.011	-0.62	.049	-.011	-0.59
$\gamma_{11} = 0.00$	.081	.000	0.01	.048	-.010	-0.59	.049	-.011	-0.56
$\gamma_{11} = 0.20$	.081	.000	0.01	.048	-.010	-0.59	.049	-.010	-0.56

Note. NJ = Average number of cases per group. NG = Number of groups. Standard errors are averages based on the data generating model. Bias was calculated by subtracting the standard error for the data generating model from the misspecified model.

Misspecification of the random part of the model did influence the standard errors for  $\gamma_{10}$  and  $\gamma_{11}$ . Cohen's *d*'s ranged from -1.14 to 0.09 for  $\gamma_{10}$  and -1.08 to 0.09 for  $\gamma_{11}$ . The smallest values were found for the no slope variance condition, and the biggest values were found for the largest group size condition. When there was no population slope variance, estimating a model with a random slope produced a bias of .001 (Cohen's *d* = 0.09) in the standard errors for both  $\gamma_{10}$  and  $\gamma_{11}$ . In contrast, misspecifying the slope as fixed when it should have been random produced biases ranging from -0.019 to -0.005 for both  $\gamma_{10}$  and  $\gamma_{11}$ . Cohen's *d* ranged from -1.03 to -0.28 for  $\gamma_{10}$  and -0.96 to -0.26 for  $\gamma_{11}$ . As expected, bigger biases occurred for greater levels of slope variance.

## ANOVA Results

To get a better sense of the effects of the study conditions on bias, we conducted separate ANOVA's investigating the impact of the study characteristics on the bias in standard errors for the  $\gamma_{10}$  and  $\gamma_{11}$ . We included only the conditions where there was slope variance to produce a fully crossed design. Table 2 presents the results. Results were generally consistent with those for both ANOVA's. Not surprisingly, the biggest effect



size was for the amount of slope variance (partial  $\eta^2 = .65$  for  $\gamma_{10}$  and  $.62$  for  $\gamma_{11}$ ). The number of groups (partial  $\eta^2 = .38$  for  $\gamma_{10}$  and  $.36$  for  $\gamma_{11}$ ) also had a sizable effect size. The partial  $\eta^2$  for group size was  $.11$  for  $\gamma_{10}$  and  $.10$  for  $\gamma_{11}$ . Based on Table 1, it appears that both larger numbers of groups and group size produced less bias. Note that as the number of groups increased, the pooled variance of bias decreased. Thus, the Cohen's  $d$  for bias as a function of the number of groups did not exhibit a similar pattern as the raw bias. The slope mean and the amount of intercept – slope correlation did not have large effect sizes (partial  $\eta^2$ s =  $.00$ ). Finally, only the interaction between the number of groups and amount of slope variance had a sizeable effect (partial  $\eta^2$ s =  $.11$ ). Figure 1 displays the form of the interaction for  $\gamma_{10}$ . The effect for slope variance decreased as the number of groups increased.

**Table 2**

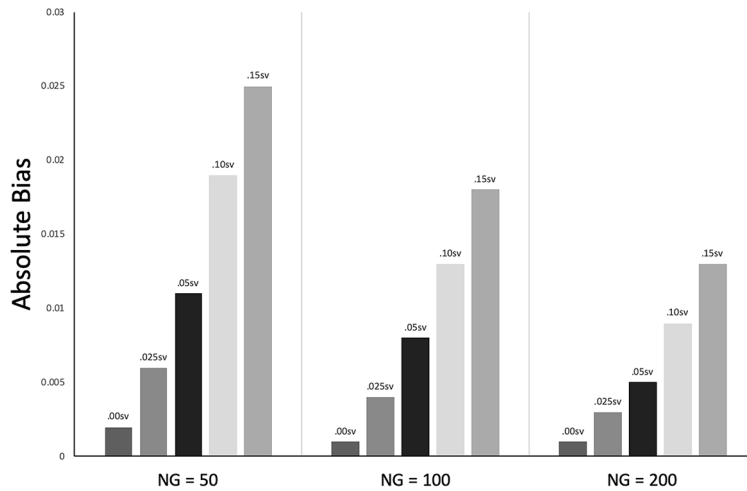
*Between Subjects ANOVA for Bias in the Standard Errors for  $\gamma_{10}$  and  $\gamma_{11}$  From Study 1*

Source	df	Bias in $\gamma_{10}$ SE		Bias in $\gamma_{11}$ SE	
		F	Partial $\eta^2$	F	Partial $\eta^2$
Number of cases per group (NJ)	2	54013.35*	.11	50272.89*	.10
Number of groups (NG)	2	264087.34*	.38	240527.90*	.36
Slope Mean (SM)	3	0.54	.00	0.66	.00
Slope variance (SV)	3	521630.03*	.65	466481.81*	.62
Cross-level interaction (CINT)	1	2.51	.00	3.86*	.00
Intercept – slope correlation	2	660.27*	.00	576.28*	.00
NG $\times$ SV	6	18279.90*	.11	16739.40	.10

*Note.* All other interactions explained less than 10 percent of the variance.

\* $p < .05$ .

Figure 1



Absolute bias of the standard errors for  $\gamma_{10}$  as a function of the amount of slope variance and number of groups.

## Study 2

In Study 2, we assessed the effects of misspecifying the random part of the model when there were multiple random slopes by adding a second Level 1 predictor. There is potential for greater effects as this also influences the correlations between that slope variance with the variances of the intercept and other slopes. We were interested in how misspecifying slope variance for  $X_1$  impacted the standard errors for the slope mean for  $X_2$  ( $\gamma_{20}$ ) and how misspecifying slope variance for  $X_2$  affected the standard errors for the slope mean for  $X_1$  ( $\gamma_{10}$ ) and the cross-level interaction ( $\gamma_{11}$ ).

## Method

### Design

We used a 3 (number of groups)  $\times$  3 (average group size)  $\times$  4 (slope correlation)  $\times$  5 (slope variance effect size for  $X_1$ )  $\times$  2 (slope variance effect size for  $X_2$ )  $\times$  4 (predictor correlation)  $\times$  2 (between group variance for  $X_1$ )  $\times$  2 (between group variance for  $X_2$ ) design. It was not fully crossed because there can be no slope correlation when either of the slopes have zero variance. The values for the number of groups, average group size, and slope variance effect size for  $X_1$  were consistent with Study 1. The cross-level interaction and slope mean of  $X_1$  were set at 0.20 and 0.30 respectively. We eliminated the other conditions for these variables because Study 1 demonstrated negligible differences in standard error bias across these conditions. In addition, we added another Level 1

predictor ( $X_2$ ). We set the correlation between the two Level 1 predictors ( $r_{X_1X_2}$ ) to be 0.00, 0.50, and 0.90. The slope mean for  $X_2$  was also set to 0.30, and the slope variance was set to either 0.00 or 0.10. We set the correlation between the two slopes to be .00, .50, or .90. When the slope variance for  $X_1$  or  $X_2$  was 0.00, the correlation between slopes was fixed to zero. Since Level 1 predictors can also contain between group variance, we set the ICC for each predictor to be .00 or .50. The latter value represents a substantial ICC value. For each possible combination, we simulated 1000 samples resulting in 3,168,000 samples.

### Data Generation

The Level 2 variables  $Z$  and  $u_{0j}$  were generated as in Study 1. For  $u_{1j}$ , the standard deviation was set to the value to produce the desired slope variance. For  $u_{2j}$ , the standard deviation was set to the value of either 0.00 or the square root of 0.10 depending on the condition. We transformed these values to produce the desired covariance matrix for random effects using procedures adapted from nonparametric bootstrapping (Goldstein, 2011). For  $X_1$  and  $X_2$ , we randomly sampled values from a standard normal distribution to produce the desired within group variance. Between group variance for  $X_1$  and  $X_2$  was created by adding the desired between group variances to the  $X_1$  and  $X_2$  predictors. In addition, the Level 1 error term,  $r_{ij}$  was created as in Study 1. Values for  $Y$  were generated by multiplying the coefficients by the relevant predictors and adding the error terms.

### Analysis

As in Study 1, we estimated multiple models. In Model 1, the slope for  $X_1$  was fixed and the slope for  $X_2$  was random. In Model 2, the  $X_1$  slope was random and the  $X_2$  slope was fixed. Finally, Model 3 specified both slopes as random. We compared the Model 1 standard errors for the fixed effects against those from Model 3 to determine the effects of misspecifying the slope of  $X_1$  on the standard errors of  $\gamma_{01}$ ,  $\gamma_{10}$ ,  $\gamma_{11}$ , and  $\gamma_{20}$ . We also compared Model 2 with Model 3 to determine if misspecifying the slope for  $X_2$  affected the standard errors for  $\gamma_{11}$ .

## Results

### Comparing Model 1 to Model 3

Table 3 displays the bias in standard errors as a function of comparing Model 1 with Model 3. Overall, the results for the standard errors for  $\gamma_{01}$ ,  $\gamma_{10}$ , and  $\gamma_{11}$  were similar to the one-predictor case. Misspecifying the slope of  $X_1$  had little effect on the standard errors of  $\gamma_{01}$  with bias estimates ranging from .000 to .002 and Cohen's  $d$ 's ranging from 0.01 to 0.07. There was considerable bias in the standard errors for both  $\gamma_{10}$  and  $\gamma_{11}$  when the slope of  $X_1$  was incorrectly fixed to zero. Bias estimates ranged from -.016 to -.005 with Cohen's  $d$ 's ranging from -0.58 to -0.16 for the  $\gamma_{10}$  standard errors. Similarly, bias

estimates for the  $\gamma_{11}$  standard errors ranged from -.004 to -.016 with Cohen's  $d$ 's ranging from -.070 to -.021. As with Study 1, the effects of misspecifying the  $\tau_{11}$  as random when it should have been fixed to zero were not as pronounced. Bias estimates for standard errors were .002 for  $\gamma_{10}$  and .001 for  $\gamma_{11}$ . Cohen's  $d$ 's were 0.08 and 0.07, respectively.

**Table 3**

*Bias Statistics for the Standard Error Differences in the Parameter Estimates Between a Model With a Fixed Slope and a Model With a Slope Allowed to Vary*

Source	$\gamma_{01}$			$\gamma_{10}$			$\gamma_{11}$			$\gamma_{20}$		
	SE	SE Bias	$d$	SE	SE Bias	$d$	SE	SE Bias	$d$	SE	SE Bias	$d$
NJ = 5	.076	.001	0.03	.082	-.010	-0.26	.067	-.008	-0.36	.081	.002	0.06
NJ = 10	.064	.001	0.03	.060	-.009	-0.35	.050	-.008	-0.51	.059	.003	0.10
NJ = 20	.059	.001	0.03	.048	-.010	-0.47	.042	-.010	-0.63	.049	.004	0.18
NG = 50	.090	.001	0.06	.088	-.013	-0.38	.074	-.012	-0.59	.087	.004	0.12
NG = 100	.064	.001	0.07	.060	-.009	-0.36	.050	-.008	-0.60	.060	.003	0.11
NG = 200	.045	.000	0.07	.042	-.006	-0.35	.035	-.006	-0.62	.042	.002	0.12
$\tau_{11} = .000$	.066	.000	-0.01	.051	.002	0.08	.040	.001	0.07	.058	.001	0.02
$\tau_{11} = .025$	.066	.000	0.01	.058	-.005	-0.16	.047	-.004	-0.21	.063	.001	0.04
$\tau_{11} = .050$	.066	.000	0.02	.061	-.008	-0.24	.050	-.007	-0.34	.063	.002	0.07
$\tau_{11} = .10$	.067	.001	0.04	.067	-.012	-0.38	.057	-.011	-0.54	.063	.004	0.13
$\tau_{11} = .15$	.067	.002	0.06	.071	-.016	-0.51	.062	-.016	-0.70	.064	.006	0.18
$\tau_{22} = .00$	.066	.001	0.03	.061	-.010	-0.28	.048	-.008	-0.40	.053	.001	0.03
$\tau_{22} = .10$	.067	.001	0.03	.064	-.010	-0.30	.054	-.009	-0.42	.066	.004	0.11
$\tau_{12} = .00$	.066	.001	0.03	.062	-.009	-0.28	.050	-.008	-0.38	.059	.001	0.04
$\tau_{12} = .10$	.067	.001	0.03	.065	-.010	-0.32	.054	-.009	-0.45	.066	.003	0.08
$\tau_{12} = .50$	.067	.001	0.03	.065	-.010	-0.31	.056	-.010	-0.46	.066	.005	0.14
$\tau_{12} = .90$	.067	.001	0.04	.065	-.010	-0.32	.056	-.010	-0.42	.066	.007	0.19
$ICC_{X_1} = .00$	.066	.000	0.02	.062	-.010	-0.30	.052	-.009	-0.43	.063	.003	0.10
$ICC_{X_1} = .50$	.067	.001	0.04	.064	-.009	-0.29	.054	-.009	-0.40	.063	.003	0.09
$ICC_{X_2} = .00$	.066	.001	0.03	.063	-.010	-0.31	.053	-.009	-0.42	.061	.003	0.10
$ICC_{X_2} = .50$	.067	.001	0.03	.064	-.010	-0.29	.053	-.009	-0.41	.064	.003	0.09
$r_{X_1X_2} = .00$	.068	.001	0.03	.052	-.010	-0.58	.051	-.010	-0.49	.052	.001	0.06
$r_{X_1X_2} = .10$	.067	.001	0.03	.056	-.010	-0.40	.051	-.009	-0.45	.056	.002	0.07
$r_{X_1X_2} = .50$	.066	.001	0.03	.056	-.010	-0.47	.052	-.010	-0.50	.055	.005	0.21
$r_{X_1X_2} = .90$	.065	.001	0.03	.090	-.009	-0.20	.058	-.007	-0.27	.089	.005	0.10

*Note.* NJ = Average number of cases per group. NG = Number of groups. ICC = Intraclass correlation coefficient. Bias in  $\gamma_{01}$ ,  $\gamma_{10}$ ,  $\gamma_{11}$ , and  $\gamma_{20}$  standard errors were calculated by subtracting those from those from a model with a random  $X_1$  slope (Model 2) from these standard errors from a model with a fixed  $X_1$  slope (Model 1).

It is interesting to note that varying the amount of slope variance produced little bias in the standard errors for  $\gamma_{20}$ . In general, the amount of bias was similar to that for the standard errors of  $\gamma_{01}$ . In addition, the degree of correlation between slopes did not appear to impact the bias in the standard errors for  $\gamma_{10}$  and  $\gamma_{11}$ . For example, when the correlation was .00, the bias in the standard errors for  $\gamma_{10}$  was -.009 with a Cohen's  $d$  of -0.28. When the correlation equaled .90, the bias equaled -.010 with a Cohen's  $d$  of -0.32. The amount of between group variance in the Level 1 predictors had little effect on the bias in the standard errors. Bias estimates for the standard errors of  $\gamma_{10}$  were -.010 with a Cohen's  $d$  of -0.30 when there was no between group variance in  $X_1$ , and -.009 with a Cohen's  $d$  of -0.29 when between group variance was 0.5. We found similar results for  $\gamma_{11}$ . The between group variance conditions for  $X_2$  produced biases and effect sizes that mimicked the findings for  $X_1$ .

We conducted two ANOVA's investigating the impact of the study characteristics on the bias in standard errors for the average  $\gamma_{10}$  and  $\gamma_{11}$ . As with Study 1, only conditions where there was variance in the slopes were included. We limited the ANOVA to the main effects and all two-way interactions because of the large number of conditions. Table 4 presents the results. Results were generally consistent with those for both ANOVA's. The largest effects were associated with the amount of slope variance (partial  $\eta^2 = .43$  for  $\gamma_{10}$  and .49 for  $\gamma_{11}$ ) and the number of groups (partial  $\eta^2 = .27$  for both  $\gamma_{10}$  and  $\gamma_{11}$ ). None of the two-way interactions a partial  $\eta^2$ s greater than .06.

**Table 4**

*Between Subjects ANOVA for Bias in the Standard Errors for  $\gamma_{10}$  and  $\gamma_{11}$ , From Study 2*

Source	df	Bias in $\gamma_{10}$ SE		Bias in $\gamma_{11}$ SE	
		F	Partial $\eta^2$	F	Partial $\eta^2$
Number of cases per group	2	9961.09	0.01	50982.95	0.04
Number of groups	2	427913.47	0.27	436516.19	0.27
Predictor correlation	3	23328.24	0.03	91597.76	0.11
Slopes correlation	3	121.52	0.00	3532.56	0.00
Slope variance for X1	3	573301.99	0.43	745638.19	0.49
ICC for X1	1	1746.76	0.00	2459.09	0.00
ICC for X2	1	118.75	0.00	299.74	0.00

Note. All two-way interactions were tested and resulted in Partial  $\eta^2 < .02$ . ICC = Intraclass correlation coefficient.

### Comparing Model 2 to Model 3

Varying the slope variance for  $X_2$  ( $\tau_{22}$ ) when it should have been fixed did not affect the standard errors for  $\gamma_{10}$  or  $\gamma_{11}$ . The bias estimates were -.001 and .000 with a Cohen's  $d$  of -0.03 and 0.01 for the standard errors for  $\gamma_{10}$  and  $\gamma_{11}$ , respectively. Similarly, fixing the

slope variance for  $X_2$  ( $\tau_{22}$ ) when it should have been random also did not have a large effect. The bias estimates were .004 and .003 with a Cohen's  $d$  of 0.11 and 0.12 for  $\gamma_{10}$  and  $\gamma_{11}$ , respectively. This provides further evidence that the effects of misspecifying the random part of multilevel models are isolated to the standard errors for the fixed effects associated with the misspecified slope variance.

## Discussion

It is well known that misspecification of the random part of multilevel models impacts the standard errors for the fixed effects. In the present studies, we sought to identify how great this impact is, and what type of misspecification has the larger effect. In Study 1, we found that fixing the slope variance when it should be random has bigger effects than freeing the slope variance when it should be fixed. Large effect sizes for both the Level 1 slope and cross-level interaction standard errors were observed for relatively small levels of slope variance. The bias was such that Type I errors were more likely when the slope was fixed but should have varied, and Type II errors were more likely when the slope varied but should have been fixed. We also found that the number of groups attenuated the biasing effect of slope variance magnitude such that more groups were associated with less of a biasing effect.

In Study 2, we investigated the effects of misspecifying the random part of the model in situations with multiple random slopes and between-group variance for correlated Level 1 predictors. Overall, the results were similar to the single random slope models in that these added conditions did not have a large impact. As expected, the bias in standard errors is limited to the fixed effects associated with the slope that should be random. That is, the only fixed effects that are affected are the direct effect of the Level 1 predictor, and the cross-level interaction of the Level 2 variable predicting the Level 1 slope.

## Limitation and Future Research

In our simulations, we considered a basic two-level model with relatively few continuous predictors. Thus, it is not clear how well the results would generalize to studies with a larger number of predictors, additional levels, or categorical variables. Similarly, we considered a model with a simple compound symmetry variance/covariance matrix for the Level 1 errors. The assumption of compound symmetry is often relaxed for longitudinal models because of autocorrelation and heterogeneity of variance that occurs for repeated measures. This directly affects the standard errors for Level 1 coefficients (Kwok, West, & Green, 2007), but it is unclear as to how this may interact with misspecifying random slopes.

## Recommendations

We see three major recommendations emerging from the results of the present studies. First, we suggest that researchers, when in doubt, give preference to random slopes. Our results suggest that doing so would minimize the amount of bias in the relevant standard errors for the fixed effects. We suspect that it is rare that the variance for a slope is exactly zero. However, even if this were to be the case, our results suggest that allowing that slope to vary would produce very little bias. In contrast, fixing a slope that should be random produces considerably more bias. If researchers desire to gauge the potential effects of misspecifying slope variance, then they could compare models with and without random slopes.

Second, we suggest that researchers should avoid using nonrandomly varying slope models to evaluate cross-level interactions. Cross-level interactions may exist when the variance in slopes is not statistically significant (LaHuis & Ferguson, 2009; Mathieu et al., 2012). Thus, researchers are advised to test hypothesized cross-level interactions even when the predicted slope does not exhibit significant slope variance. What is not clear is whether the relevant slope variance should be fixed to zero. In the present studies, we found large effect sizes for bias in the standard errors for cross-level interactions when this was done and may lead to Type I errors for cross-level interactions.

Finally, we recommend using a piecemeal approach when there are many random effects that may cause estimation problems or if there are not enough degrees of freedom to estimate all of the random effects. Results from Study 2 suggest that omitting a slope from the random part of the model only affects coefficients concerning the omitted slope. Researchers could test their hypotheses regarding fixed effects allowing for the appropriate random effects. For example, we may hypothesize that the Level 2 variable,  $Z$ , predicts the slopes of  $X_1$  and  $X_2$ . We could first examine the effects of  $Z$  on the  $X_1$  slope while allowing the  $X_1$  slope to vary and fixing the  $X_2$  slope to zero. Next, we would examine the effects of  $Z$  on the  $X_2$  slope while allowing the  $X_2$  slope to vary and fixing the  $X_1$  slope. Doing so would allow us to examine the hypotheses using the appropriate standard errors.

---

**Funding:** The authors have no funding to report.

---

**Competing Interests:** The authors have declared that no competing interests exist.

---

**Acknowledgments:** The authors have no support to report.

---

## References

- Baldwin, S. A., Bauer, D. J., Stice, E., & Rohde, P. (2011). Evaluating models for partially clustered designs. *Psychological Methods, 16*(2), 149-165. <https://doi.org/10.1037/a0023464>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). <https://doi.org/10.18637/jss.v067.i01>
- Berkhof, J., & Kampen, J. K. (2004). Asymptotic effect of misspecification in the random part of the multilevel model. *Journal of Educational and Behavioral Statistics, 29*(2), 201-218. <https://doi.org/10.3102/10769986029002201>
- Chung, H., & Beretvas, S. N. (2012). The impact of ignoring multiple membership data structures in multilevel models. *British Journal of Mathematical & Statistical Psychology, 65*(2), 185-200. <https://doi.org/10.1111/j.2044-8317.2011.02023.x>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.
- Gilbert, J., Petscher, Y., Compton, D. L., & Schatschneider, C. (2016). Consequences of misspecifying levels of variance in cross-classified longitudinal data structures. *Frontiers in Psychology, 7*, Article 695. <https://doi.org/10.3389/fpsyg.2016.00695>
- Goldstein, H. (2011). Bootstrapping in multilevel models. In J. J. Hox & J. K. Roberts (Eds.), *Handbook of advanced multilevel analysis* (pp. 163-171). New York, NY, USA: Taylor Francis Group.
- Hox, J., Moerbeek, M., & van de Schoot, R. (2018) *Multilevel analysis: Techniques and applications* (3rd ed.). Mahwah, NJ, USA: Lawrence Erlbaum Associates.
- Korendijk, E. J. H., Hox, J. J., Moerbeek, M., & Maas, C. J. M. (2011). Robustness of parameter and standard error estimates against ignoring a contextual effect of a subject-level covariate in cluster-randomized trials. *Behavior Research Methods, 43*(4), 1003-1013. <https://doi.org/10.3758/s13428-011-0094-8>
- Korendijk, E. J. H., Maas, C. J. M., Moerbeek, M., & Van der Heijden, P. G. M. (2008). The influence of misspecification of the heteroscedasticity on multilevel regression parameter and standard error estimates. *Methodology, 4*(2), 67-72. <https://doi.org/10.1027/1614-2241.4.2.67>
- Kwok, O., West, S. G., & Green, S. B. (2007). The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A Monte Carlo study. *Multivariate Behavioral Research, 42*(3), 557-592. <https://doi.org/10.1080/00273170701540537>
- LaHuis, D. M., & Ferguson, M. W. (2009). The accuracy of statistical tests for variance components in multilevel random coefficient modeling. *Organizational Research Methods, 12*(3), 418-435. <https://doi.org/10.1177/1094428107308984>
- LaHuis, D. M., Hartman, M. J., Hakoyama, S., & Clark, P. C. (2014). Explained variance measures for multilevel models. *Organizational Research Methods, 17*(4), 433-451. <https://doi.org/10.1177/1094428114541701>
- Luo, W., & Kwok, O.-M. (2009). The impacts of ignoring a crossed factor in analyzing cross-classified data. *Multivariate Behavioral Research, 44*(2), 182-212. <https://doi.org/10.1080/00273170902794214>



- 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. *The Journal of Applied Psychology*, 97, 951-966. <https://doi.org/10.1037/a0028380>
- Meyers, J. L., & Beretvas, S. N. (2006). The impact of inappropriate modeling of cross-classified data structures. *Multivariate Behavioral Research*, 41(4), 473-497. [https://doi.org/10.1207/s15327906mbr4104\\_3](https://doi.org/10.1207/s15327906mbr4104_3)
- Moerbeek, M. (2004). The consequence of ignoring a level of nesting in multilevel analysis. *Multivariate Behavioral Research*, 39(1), 129-149. [https://doi.org/10.1207/s15327906mbr3901\\_5](https://doi.org/10.1207/s15327906mbr3901_5)
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA, USA: Sage Publications.
- Raudenbush, S. W., & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. *Psychological Methods*, 5(2), 199-213. <https://doi.org/10.1037/1082-989X.5.2.199>
- R Core Team. (2018). *R: A language and environment for statistical computing* [R Foundation for Statistical Computing, Vienna, Austria]. Retrieved from <https://www.R-project.org/>
- Shi, Y., Leite, W., & Algina, J. (2010). The impact of omitting the interaction between crossed factors in cross-classified random effects modelling. *British Journal of Mathematical & Statistical Psychology*, 63(1), 1-15. <https://doi.org/10.1348/000711008X398968>
- Snijders, T. A. B., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Thousand Oaks, CA, USA: Sage Publications.



*Methodology* is the official journal of the European Association of Methodology (EAM).



[leibniz-psychology.org](https://leibniz-psychology.org)

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology Information (ZPID), Germany.