Conflated Random Slopes in Multilevel Analysis: A Comprehensive Review

1. Introduction

1.1 An Overview of Multilevel Analysis

Multilevel analysis, also known as hierarchical linear modeling or linear mixed-effects modeling, has become a cornerstone method in various fields including education, sociology, and psychology. Its popularity stems from its ability to account for the nested structure of clustered data (e.g., students at level 1 within schools at level 2), in which the dependence between individual units of the same cluster poses a substantial challenge for traditional methods based on ordinary least squares estimation (Gelman & Hill, 2007; Finch et al., 2019).

Multilevel analysis enables researchers to decompose the variance of individual outcomes according to within- and between-cluster differences, which serve to examine intricate relationships between individual and cluster-level characteristics (Raudenbush & Bryk, 2002). The usefulness of this method lies in its ability to handle complex data dependencies, provide accurate standard errors, and offer insights into how contextual factors predict individual outcomes.

1.2 Research Problem: Conflated Random Slopes

Over the past two decades, methodologists have identified several challenges in the application of multilevel analysis. One particularly troublesome issue that has received increasing attention is the problem of *conflated random slopes* (Rights & Sterba, 2023). This occurs when the slope heterogeneity and the intercept heteroscedasticity of the within-cluster and the between-cluster parts of an individual predictor are conflated or smushed into one estimate in the model. This leads to biased parameter estimates and potentially misleading conclusions (Rights, 2022). To avoid conflated random slopes, level 1 predictors must be adequately parameterized.

This literature review examines the development of research on conflated random slopes in multilevel modeling over the last two decades. It synthesizes theoretical perspectives, methodological advancements, and empirical applications to provide a comprehensive understanding of the problem and its solutions. This review looks for evidence to answer three research questions: what is the origin of the conflated random slopes? What are the consequences of these slopes for multilevel analysis? Finally, what methodological decisions help to avoid the estimation of these slopes? The review is structured to address the background of the problem, its consequences for research, methodological solutions, and directions for future research.

2. Theoretical Background

2.1 Basic Components of Multilevel Analysis

Multilevel analysis, as developed from the combination of contextual analysis and the statistical theory of mixed models in the 1980s (De Leeuw et al., 2008), emerged as a solution to the analytical challenges posed by clustered data, which basic premise is that individuals within the same cluster are more similar to each other than observations from different clusters, violating the independence assumption of traditional regression models (Hamaker & Muthén, 2020). Multilevel models (MLMs) address this issue by incorporating random effects in the model, which introduce specific patterns of heteroscedasticity and correlation in the variance-covariance matrix of individuals from the same cluster. Additionally, random effects have variances that break down the total outcome variance and allow researchers to analyze the predictive capacity of individual and cluster predictors.

A multilevel model (MLM) consists of fixed and random effects that capture relationships between variables in nested data, typically featuring two primary levels: level 1 for individual observations and level 2 for clusters. Each level works as a submodel (McNeish, 2023). Level 1 predicts the individual outcome as a function of the fixed effects of individual predictors, whereas level 2 includes fixed effects for cluster attributes and random effects for level 1 predictors. The centering techniques are crucial for the model specifications, as they help to disaggregate the within- and between-cluster effects, enabling more precise estimation of fixed and random effects while maintaining the model's ability to account for the hierarchical structure of the data (more on this later).

The basic random effect of any multilevel model is the random intercept, which captures differences between specific cluster mean outcomes and the overall outcome mean predicted by the model. The random intercept allows for the quantification and prediction of constant mean differences between clusters, establishing the structure of the multilevel model: individual differences within clusters at level 1 and differences between clusters at level 2 (Snijders & Bosker, 2012). Then, predictors can be added at level 1 and level 2 to explain these variances, and random slopes can be added to capture potential differences in the fixed effect of an individual predictor across clusters at level 2 (which makes multilevel analysis very attractive for evaluating the relevance of the context of individuals). In addition, cross-level interactions can be included to analyze how the effect of individual predictors is moderated by context attributes (Loeys, Josephy, & Dewitte, 2018).

This summary is sufficient to understand the main discussion of this document on conflated random slopes. For readers interested in more details on how to apply multilevel analysis, we recommend reading Raudenbush and Bryk (2002) and Snijders and Bosker (2012). Next, we will review three aspects of multilevel analysis that are directly related to the problem of conflated random slopes: the relevance of level 1 predictors in multilevel analysis, the role of centering techniques in the disaggregation of level-specific effects, and the problem of conflated fixed effects.

2.2 Relevance of Level 1 Predictors

In multilevel analysis, level 1 predictors require special attention, given that their relationship with the outcome can differ across levels (Rights et al., 2020; Snijders, 2005). This is because level 1 predictors may have individual and cluster information that must be disaggregated to avoid biases in the estimation and interpretation of fixed effects (Curran et al., 2012; Preacher et al., 2010). A common solution to this problem is to add the cluster means of the level 1 predictor to the model, so both individual and cluster predictors get separate fixed effect (Enders, 2013). To illustrate, let us consider an example scenario of students (as individuals) nested in classrooms (as clusters), as shown in Equation 1:

Level 1

$$y_{sc} = \beta_{0c} + \beta_{1c}(x_{sc}) + e_{sc}$$

$$e_{sc} \sim N(0, \sigma_e^2),$$

Level 2

$$\beta_{0c} = \gamma_{00} + \gamma_{01}^{c}(x_{.c}) + U_{0c}$$

$$\beta_{1c} = \gamma_{10}$$

$$U_{0c} \sim N\left(0, \tau_{U_0}^2\right), \tag{1}$$

in which the individual outcome y_{sc} (e.g., a test score) for the student s in classroom c is being predicted by an individual predictor x_{sc} (e.g., student GPA centered at 3) and its cluster mean $x_{.c}$ (i.e., classroom-mean GPA centered at 3). The β terms at level 1 are placeholders (or latent variables) composed of fixed and random effects at level 2 (Snijders & Bosker, 2012). Specifically, β_{0c} is the predicted cluster-specific intercept composed of the fixed intercept γ_{00}, i.e., the conditional outcome mean across clusters when all predictors are zero, the contextual fixed effect γ_{01}^{c} for $x_{.c}$, plus U_{0c}, a specific cluster-deviation from γ_{00}. The contextual fixed effect of the classroom GPA captures the incremental contribution of this classroom attribute on the prediction of the test scores after controlling for the student's GPA.

The term β_{1c} in Equation 1 is the predicted cluster-specific slope for student GPA, composed of the within-cluster fixed effect γ_{10}. The within-cluster fixed effect captures the expected change in the predicted outcome per unit difference in student GPA within a classroom c. Finally, the individual-specific and cluster-specific residuals, e_{sc} and U_{0c}, are both assumed to be independent and normally distributed random variables with mean 0 and estimated variances σ_e^2 and $\tau_{U_0}^2$, respectively. Both level-specific residuals are not related to each other nor to the predictors of the other level (Goldstein, 2011; Shaw & Flake, 2024).

2.3 Centering Techniques

The regression coefficient γ_{01} in Equation 1 would be interpreted as a between-cluster fixed effect (labeled as γ_{01}^{b}) if the level 1 predictor had been centered on its respective cluster means (i.e., $x_{sc} - x_{.c}$). A between-cluster fixed effect captures the expected change in the outcome per unit difference in the level 2 predictor across clusters. The contextual or between-cluster fixed effects depend on the centering of the level 1 predictor: *constant-centering* or CC (Hoffman & Walters, 2022) or *centering-within-cluster* or CWC (Enders & Tofighi, 2007).

When CC is used, as in Equation 1, the relationship between predictors at different levels is not orthogonal, as the original predictor retains the same information as before. Therefore, the level 2 cluster means are necessary to disaggregate the within-cluster and contextual between-cluster fixed effects

(Hoffman & Walters, 2022). The contextual fixed effects also serve to test whether the between-cluster fixed effect is distinct from the within-cluster fixed effect of a level 1 predictor, given that this coefficient is equal to the difference between the former and the latter (Raudenbush & Bryk, 2002). When CWC is used, the relationship between the predictors is orthogonal, which creates purely level-specific (within- and between-cluster) fixed effects, so the cluster means are not required in the model—although it is advisable to maintain them.

Centering is directly related to the problem of conflated fixed and random effects in multilevel analysis. As a level 1 predictor may contain information about the clusters, by using only this predictor in the model we would be forcing its fixed effect to conflate intra- and inter-cluster differences in the prediction of the outcome (Hamaker & Muthén, 2020).

2.4 Conflated Fixed Effects

The problem of conflated fixed effects in multilevel analysis emerges from the complex nature of level 1 predictors, which can simultaneously contain both within- and between-cluster information. Originally discussed by methodologists in the late 1980s and early 2000s (including pioneer work by Cronbach & Webb, 1975; Raudenbush, 1989; Longford, 1989; and Raudenbush & Bryk, 2002), conflated fixed effects distort the understanding of the relationship of the level 2 component within the level 1 predictor with the level 2 outcomes and lead to biased statistical inferences (Hamaker & Muthén, 2020; Rights & Sterba, 2016; Rights et al., 2020).

Conflated fixed effects, also known as smushed, blended, or composite fixed effects in the multilevel analysis literature, are an uninterpretable weighted average of the within- and between-cluster fixed effects attributable to the individual and cluster information found in the level 1 predictors, respectively (Hedeker & Gibbons, 2006; Hoffman, 2015; Hoffman & Walters, 2022; McNeish, 2023; Preacher et al., 2010; Preacher, 2011). Substantially, researchers interpreting these fixed effects may misunderstand the true relationship between predictors and outcomes and draw erroneous conclusions about cluster-level variations.

Methodological solutions have predominantly focused on centering techniques, particularly constant-centering (CC) and centering-within-cluster (CWC), which help separate within-cluster and between-cluster effects (Hoffman & Walters, 2022). More recent advances, such as those proposed by Preacher et al. (2016) and Rights and Sterba (2023), have developed more nuanced multilevel models based on Structural Equation Modeling (SEM) and level-specific parameterization of fixed and random effects, respectively, which can directly address and mitigate the problem of effect conflation.

3. Conflated Random Slopes

3.1 Definition of the Problem

The problems of conflated fixed effects and how to avoid them have been studied extensively in the multilevel analysis literature (Hoffman, 2019; Rights, 2022; Snijders & Bosker, 2012). However, the problem of conflated random slopes has not been covered with the same intensity (but see Rights & Sterba, 2023 for an exception). Although some papers from the 1980s had mentioned the differences in the interpretation of the random slopes depending on the centering of the level 1 predictor (Longford, 1989; Raudenbush, 1989; Plewis, 1989), their focus was largely on the interpretation of the model intercept.

Twenty years passed since those discussions, and the works of Lüdtke et al. (2008) and Preacher et al. (2010) once again mentioned the importance of separating the within- and between-cluster components of individual predictors when estimating random slopes. This was related to the theoretical and methodological progress of multilevel analysis, which since 2010 has expanded rapidly in the social sciences, economics and health, with several books and papers contributing to the understanding and application of this method in these areas (Hox & Roberts, 2011).

A random slope represents the heterogeneity of the fixed effect of a level 1 predictor across clusters in the model (Hox et al, 2017). However, when the level-specific parts of a level 1 predictor are not disaggregated, their fixed effects are conflated. Random slopes are not an exception, although their interpretation is different. A conflated or smushed random slope is an uninterpretable blend of both the slope heterogeneity and the intercept heteroscedasticity of the within- and between-cluster fixed effects of a level 1 predictor and its cluster means at level 2 (Rights & Sterba, 2023). The intercept heteroscedasticity is a consequence of the cluster information contained in the original level 1 predictor that was not disaggregated or controlled for, which is being inadvertently included in the variance of the cluster intercepts.

Conflated random slopes can bias the estimation, interpretation, and inferences of the level 2 variances in the MLM, as well as the calculation of the standard errors of the fixed effects and their subsequent Type I error rates (LaHuis et al., 2020; Rights et al., 2020; Rights & Sterba, 2023). When only the CC version of the level 1 predictor is used to estimate its random slope, its within- and between-cluster parts are forced to have the same variability across clusters. To understand this constraint, we can go back to the model in Equation 1 and replace the individual predictor x_{sc} with an equivalent version that reflects its duality (within- and between-cluster information): $x\_sc = x_{sc} - x_{.c} + x_{.c}$, in which $x_{sc} - x_{.c}$ corresponds to the purely within-cluster part of x_{sc}, and $x_{.c}$ corresponds to its purely between-cluster part. Then, a random slope for x_{sc} can be included, labeled as U_{1c}^{s} with the superscript s to denote its smushed parameterization. This model, known as the *conventional random-slope contextual effect MLM* (Rights & Sterba, 2023) is shown in Equation 2:

Level 1

$$y_{sc} = \beta_{0c} + \beta_{1c}([x_{sc} - x_{.c}] + x_{.c}) + e_{sc}$$

$$e_{sc} \sim N(0, \sigma_e^2),$$

Level 2

$$\beta_{0c} = \gamma_{00} + \gamma_{01}^{c}(x_{.c}) + U_{0c}$$

$$\beta_{1c} = \gamma_{10} + U_{1c}^{s}$$

$$\begin{bmatrix} U_{0c} \\ U_{1c}^{s} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{U_0}^2 & \tau_{01} \\ \tau_{01} & \tau_{U_1^s}^2 \end{bmatrix}\right). \tag{2}$$

If we check the reduced-form expression of this model in Equation 3, we will see that the conflated random slope U_{1c}^{s} (bolded for emphasis below) is being multiplied twice in the model:

$$y_{sc} = \gamma_{00} + \gamma_{10}([x_{sc} - x_{.c}] + x_{.c}) + \gamma_{01}^{c}(x_{.c}) + \boldsymbol{U_{1c}^{s}}([x_{sc} - x_{.c}] + x_{.c}) + e_{sc}$$

$$y_{sc} = \gamma_{00} + \gamma_{10}(x_{sc} - x_{.c}) + (\gamma_{01} + \gamma_{10}^{c})x_{.c} + \boldsymbol{U_{1c}^{s}}(x_{sc} - x_{.c}) + \boldsymbol{U_{1c}^{s}}(x_{.c}) + e_{sc}. \tag{3}$$

Inadvertently, U_{1c}^{s} is forcing the variability around the within- and between-cluster fixed effects of $x\_sc - x_{.c}$ and $x_{.c}$ across clusters to be equal. Thus, the conventional random-slope contextual effect MLM conflates the heterogeneity around the fixed effect of the level 1 predictor with the heteroscedasticity of the cluster-specific intercepts across the cluster mean predictor values—the latter heterogeneity is introduced by the random effect of a purely level 2 predictor (Goldstein, 2011). This is shown in Equation 4,

$$\begin{aligned} \mathrm{var}(\beta_{0c}|x_{.c}) &= \mathrm{var}(\gamma_{00} + [\gamma_{10} + \gamma_{01}^{c}]x_{.c} + U_{0c} + U_{1c}^{s}(x_{.c})|x_{.c}) \\ &= \mathrm{var}(U_{0c} + U_{1c}^{s}(x_{.c})|x_{.c}) \\ &= \mathrm{var}(U_{0c}) + 2\mathrm{cov}(U_{1c}^{s}[x_{.c}], U_{0c}|x_{.c}) + \mathrm{var}(U_{1c}^{s}[x_{.c}]|x_{.c}) \\ &= \mathrm{var}(U_{0c}) + 2\mathrm{cov}(U_{1c}^{s}, U_{0c})x_{.c} + \mathrm{var}(U_{1c}^{s})x_{.c}^{2} \\ &= \tau_{U_0}^{2} + 2\tau_{01}^{s}(x_{.c}) + \boldsymbol{\tau_{U_1^s}^{2}}(x_{.c}^{2}), \end{aligned} \tag{4}$$

in which the variance of the cluster-specific intercepts is linked to the cluster means in $x_{.c}$.

This heteroscedastic intercept variance includes a quadratic function of $x_{.c}$ that is being multiplied by the variance of the conflated random slope U_{1c}^{s}. The quadratic function allows the distribution of the cluster intercepts to be wider at the extreme values of $x_{.c}$, although in the current MLM the degree of this heterogeneity is forced to be the same as that used to describe the heterogeneity of the outcome variance conditional on the level 1 predictor with random slope. Equation 5 elaborates on this problem:

$$\begin{aligned} \mathrm{var}(y_{sc}|x_{sc}) &= \mathrm{var}(\gamma_{00} + \gamma_{01}^{c}[x_{.c}] + \gamma_{10}[x_{sc}] + U_{0c} + U_{1c}^{s}[x_{sc}] + e_{sc}) \\ &= \mathrm{var}(U_{0c} + U_{1c}^{s}[x_{sc}] + e_{sc}) \end{aligned}$$

$$= \tau^2_{U_0} + 2\tau^s_{01}(x_{sc}) + \boldsymbol{\tau^2_{U^s_1}}(x^2_{sc}) + \sigma^2_e. \quad (5)$$

3.2 Consequences of Conflated Random Slopes

The discussion of conflated random effects and their consequences is important because the conventional random-slope contextual effect MLM is widely used in the social sciences (Brincks et al., 2016; Hox & Roberts, 2011). Conflated random slopes may bias the estimation, interpretation, and inferences of between-cluster differences, leading researchers to erroneous conclusions. Conflated random slopes may mistakenly imply a lack of slope heterogeneity, or they can yield a false presence of slope heterogeneity (Rights & Sterba, 2023). The extent of these biased inferences depends on the degree of intercept heteroscedasticity across the values of the cluster mean predictor in the population. For example, if the cluster-specific intercepts in the population do not vary across the cluster mean predictor (i.e., $\tau^2_{U^b_2} \cong 0$), even when there is heterogeneity around the within-cluster fixed effect across clusters (i.e., $\tau^2_{U_1} > 0$), the variance of the conflated $\tau^2_{U^s_1}$ will be pulled toward 0, increasing the likelihood of concluding there is no slope heterogeneity when there really is (a Type I error).

The alternative error—falsely detecting slope heterogeneity—could be made when the heterogeneity around the within-cluster fixed effect is close to 0 in the population (i.e., $\tau^2_{U_1} \cong 0$), but there is intercept heteroscedasticity across the cluster mean predictor (i.e., $\tau^2_{U^b_2} > 0$), which can move the conflated $\tau^2_{U^s_1}$ away from 0 instead. In this situation, the conflated random slope would seem to indicate that the relationship between the level 1 predictor and the outcome varies across clusters, when there is actually variation in the heterogeneity of the cluster intercepts across values of the cluster mean predictor.

3.3 How to Avoid Conflated Random Slopes

To avoid the estimation of conflated random effects for level 1 predictors, the same centering options we saw above can be used (CC and CWC), which implies adding a random slope for $x_{.c}$ at level 2. When CC is used, this random slope (labeled as U^c_{2c}, whose variance $\tau^2_{U^c_2}$ is the model parameter) captures differential intercept heteroscedasticity across the cluster means, whereas when CWC is used, such random slope (labeled as U^b_{2c}, whose variance $\tau^2_{U^b_2}$ is the model parameter) captures intercept heteroscedasticity across the cluster means (Rights & Sterba, 2023). Although the idea of adding a random slope for a level 2 predictor may seem counter-intuitive due to the lack of a higher level across which the fixed effect could vary (Preacher, 2017), its role is to disaggregate sources of variance in the model (Rights & Sterba, 2016).

A recent study by Rights and Sterba (2023) has proposed different models to avoid the estimation of conflated random slopes, which are listed in Table 1. Importantly, the parametrization of the fixed effects is the same across models—within-cluster plus contextual fixed effects—on purpose to focus on the parameterization of the random slopes.

Table 1. Multilevel Models that Avoid Conflated Random Slopes

Model 1: Within-cluster random slope

Level 1

$$y_{sc} = \beta_{0c} + \beta_{1c}(x_{sc}) + \beta_{2c}(x_{sc} - x_{.c}) + e_{sc}$$

$$e_{sc} \sim N(0, \sigma_e^2),$$

Level 2

$$\beta_{0c} = \gamma_{00} + \gamma_{01}^{c}(x_{.c}) + U_{0c}$$

$$\beta_{1c} = \gamma_{10}$$

$$\beta_{2c} = U_{3c}$$

$$\begin{bmatrix} U_{0c} \\ U_{3c} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{U_0}^2 & \tau_{03} \\ \tau_{03} & \tau_{U_3}^2 \end{bmatrix}\right)$$

Model 2: Between-cluster random slope

Level 1

$$y_{sc} = \beta_{0c} + \beta_{1c}(x_{sc}) + \beta_{2c}(x_{sc} - x_{.c}) + e_{sc}$$

$$e_{sc} \sim N(0, \sigma_e^2),$$

Level 2

$$\beta_{0c} = \gamma_{00} + \gamma_{01}^{c}(x_{.c}) + U_{0c} + U_{2c}^{b}(x_{.c})$$

$$\beta_{1c} = \gamma_{10}$$

$$\beta_{2c} = U_{3c}$$

$$\begin{bmatrix} U_{0c} \\ U_{3c} \\ U_{2c}^{b} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{U_0}^2 & \tau_{03} & \tau_{02}^{b} \\ \tau_{03} & \tau_{U_3}^2 & \tau_{12}^{b} \\ \tau_{02}^{b} & \tau_{12}^{b} & \tau_{U_2^b}^2 \end{bmatrix}\right)$$

Model 3: Within-cluster and contextual random slopes

Level 1

$$y_{sc} = \beta_{0c} + \beta_{1c}(x_{sc}) + e_{sc}$$

$$e_{sc} \sim N(0, \sigma_e^2),$$

Level 2

$$\beta_{0c} = \gamma_{00} + \gamma_{01}^{c}(x_{.c}) + U_{0c} + U_{2c}^{c}(x_{.c})$$

$$\beta_{1c} = \gamma_{10} + U_{1c}$$

$$\begin{bmatrix} U_{0c} \\ U_{1c} \\ U_{2c}^{c} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{U_0}^2 & \tau_{01} & \tau_{02}^{c} \\ \tau_{01} & \tau_{U_1}^2 & \tau_{12}^{c} \\ \tau_{02}^{c} & \tau_{12}^{c} & \tau_{U_2^c}^2 \end{bmatrix} \right)$$

Model 1 exclusively adds heterogeneity around the within-cluster fixed effect of the level 1 predictor across clusters. Although it is uncommon to include different versions of the level 1 predictor in the fixed and random parts of the model, this so-called “hybrid” parameterization disaggregates the random slope variance $\tau_{U_1^s}^2$ while maintaining a contextual interpretation of the level 2 fixed effect of the cluster mean of x_{sc}. One thing to consider is that this model assumes intercept homoscedasticity, given that a term to capture intercept heteroscedasticity across the values of the cluster mean predictor at level 2 has not been included (i.e., $\tau_{U_2^b}^2 = 0$).

Models 2 and 3 add slope heterogeneity and intercept heteroscedasticity terms. In these models, $\tau_{U_2^b}^2$ or $\tau_{U_2^c}^2$ serve to capture purely or differential heterogeneity in the variance of the cluster-specific intercepts across the values of the cluster mean predictor, respectively. It is worth remembering that when the CWC version of the level 1 predictor is used to estimate its random slope, U_{2c}^{b} captures intercept heteroscedasticity *between* clusters directly (as we are manually disaggregating both sources of heterogeneity). In contrast, when the CC version of the level 1 predictor is used to estimate its random slope, U_{2c}^{c} captures *differential* intercept heteroscedasticity (as we have then corrected the level 2 intercept heteroscedasticity from any bias due to $\tau_{U_1}^2$). A comprehensive summary of the main features of each model is included in Table 2.

Table 2. Main Features of Each Multilevel Model of Interest

Model	**Random Slopes**	**What is being added to the model?**	**Conflated effect?**	**Key assumption**
(1) Within-cluster random slope	$U_{3c}(x_{sc} - x_{.c})$	Purely slope heterogeneity	No	Intercept homoscedasticity $\tau_{U_2^b}^2 = 0$
(2) Between-cluster random slope	$U_{3c}(x_{sc} - x_{.c})$, $U_{2c}^{b}(x_{.c})$	Slope heterogeneity and intercept heteroscedasticity	No	
(3) Within-cluster and contextual	$U_{1c}(x_{sc})$,	Slope heterogeneity and differential	No	

random slopes	$U_{2c}^{c}(x_{.c})$	intercept heteroscedasticity		

4. Future Directions

Future research on conflated random slopes should address both methodological and practical challenges, including the disaggregation of effects in non-linear and non-normal models, extensions to three-level and cross-classified models, the development of open-source software, the establishment of reporting standards, and the creation of a comprehensive conceptual framework. Currently, research in this area has produced a heterogeneous body of work that, while grounded in shared ideas, lacks a consistent and comparable set of concepts across authors. Moreover, as this issue continues to evolve, its complexity remains a barrier to broader understanding among applied researchers. Therefore, future work should focus on refining the existing framework, enhancing conceptual clarity, and providing empirical evidence on the implications of conflated random slopes for parameter estimation and interpretation in multilevel analysis.

5. Conclusion

The problem of conflated random slopes has become increasingly recognized as a critical methodological issue in multilevel analysis over the past decade. What began as a specialized concern among methodologists has gained wider recognition as a fundamental challenge that can substantially impact the conclusions drawn from applied multilevel analysis across disciplines.

Regarding the initial research questions, this literature review has discussed that conflated random slopes arise when the within- and between-cluster components of an individual predictor are not disaggregated, resulting in a conflation of slope heterogeneity and intercept heteroscedasticity at level 2. Consequently, the interpretation and inferences about cluster differences may be biased, leading researchers to erroneous conclusions. To avoid conflated fixed and random slopes, centering level 1 predictors using CC or CWC is crucial. Although the interpretation of the level-specific fixed and random slopes differs between centering techniques, both serve to correctly estimate and analyze model parameters.

Future research should address both methodological and practical issues related to conflated random slopes, ensuring their accessibility to applied researchers through improved training, dissemination, and software tools. Additionally, efforts should be directed toward refining the current heterogeneous framework, which, despite sharing common ideas, lacks easily comparable concepts across authors.

References

Brincks, A., Enders, C., Llabre, M., Bulotsky-Shearer, R., Prado, G., & Feaster, D. (2017). Centering predictor variables in three-level contextual models. *Multivariate behavioral research, 52*(2), 149-163. doi:https://doi.org/10.1080/00273171.2016.1256753

Cronbach, L., & Webb, N. (1975). Between-class and within-class effects in a reported aptitude* treatment interaction: Reanalysis of a study by GL Anderson. *Journal of Educational Psychology, 67*(6), 717-724.

Curran, P., Lee, T., Howard, A., Lane, S., & MacCallum, R. (2012). Disaggregating Within-Person and Between-Person Effects in Multilevel and Structural Equation Growth Models . In J. Harring, & G. Hancock (Eds.), *Advances in Longitudinal Methods in the Social and Behavioral Sciences* (pp. 217-254). Information Age Publishing Inc.

De Leeuw, J., Meijer, E., & Goldstein, H. (2008). *Handbook of multilevel analysis.* New York: Springer.

Enders, C. (2013). Centering Predictors and Contextual Effects. In *The SAGE Handbook of Multilevel Modeling* (pp. 89-108). SAGE.

Enders, C., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods, 12*(2), 121-138. doi:10.1037/1082-989X.12.2.121

Finch, W., Bolin, J., & Kelley, K. (2019). *Multilevel modeling using R.* Chapman and Hall/CRC.

Gelman, A. H. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge university press.

Goldstein, H. (2011). *Multilevel statistical models.* John Wiley & Sons.

Hamaker, E., & Muthén, B. (2020). The fixed versus random effects debate and how it relates to centering in multilevel modeling. *Psychological Methods*, *25*(3), 365–379. doi:https://psycnet.apa.org/doi/10.1037/met0000239

Hedeker, D., & Gibbons, R. (2006). *Longitudinal data analysis.* John Wiley & Sons.

Hoffman, L. (2015). *Longitudinal analysis: Modeling within-person fluctuation and change.* Routledge.

Hoffman, L. (2019). On the Interpretation of Parameters in Multivariate Multilevel Models Across Different Combinations of Model Specification and Estimation. *Advances in Methods and Practices in Psychological Science, 2*(3), 288-311. doi:10.1177/2515245919842770

Hoffman, L., & Walters, R. (2022). Catching Up on Multilevel Modeling. *Annual Review of Psychology, 73*, 659-689. doi:https://doi.org/10.1146/annurev-psych-020821-103525

Hox, J., & Roberts, J. (Eds.). (2011). *Handbook of advanced multilevel analysis.* Psychology Press.

Hox, J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel Analysis: Techniques and Applications.* New York: Routledge.

LaHuis, D., Jenkins, D., Hartman, M., Hakoyama, S., & Clark, P. (2020). The effects of misspecifying the random part of multilevel models. *Methodology, 16*(3), 224-240. doi:https://doi.org/10.5964/meth.2799

Loeys, T., Josephy, H., & Dewitte, M. (2018). More Precise Estimation of Lower-Level Interaction Effects in Multilevel Models. *Multivariate Behavioral Research, 53*(3), 335–347. Retrieved from https://doi.org/10.1080/00273171.2018.1444975

Longford, N. (1989). To center or not to center. *Multilevel modeling newsletter*, 7-8.

Lüdtke, O., Marsh, H., 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*(3), 203-229. doi: 10.1037/a0012869

McNeish, D. (2023). A practical guide to selecting and blending approaches for clustered data: Clustered errors, multilevel models, and fixed-effect models. *Psychological Methods*, 1-27. Retrieved from https://psycnet.apa.org/doi/10.1037/met0000620

Plewis, I. (1989). Comment on "Centering" Predictors in Multilevel Analysis. *Multilevel Modeling Newsletter*, 6-7.

Preacher, K. (2011). Multilevel SEM Strategies for Evaluating Mediation in Three-Level Data. *Multivariate Behavioral Research, 46*(4), 691-731. doi:10.1080/00273171.2011.589280

Preacher, K. (2017). Multilevel Modeling: A Second Course. *Statistical Horizons*, (pp. 1-17). Philadelphia, Pennsylvania. Retrieved from https://statisticalhorizons.com/wp-content/uploads/MLM2-Sample-Materials.pdf

Preacher, K., Zhang, Z., & Zyphur, M. (2016). Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychol Methods, 21*(2), 189-205. doi:10.1037/met0000052

Preacher, K., Zyphur, M., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233. doi:10.1037/a0020141

Raudenbush, S. (1989). Centering predictors in multilevel anaysis: choices and consequences. *Multilevel modeling newsletter, 1*(2), 10-12.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods.* California: Sage Publications.

Rights, J. (2022). Aberrant Distortion of Variance Components in Multilevel Models Under Conflation of Level-Specific Effects. *Psychological Methods, 28*(5), 1154–1177. doi:https://psycnet.apa.org/doi/10.1037/met0000514

Rights, J., & Sterba, S. (2016). The relationship between multilevel models and non-parametric multilevel mixture models: Discrete approximation of intraclass correlation, random coefficient distributions, and residual heteroscedasticity. *British Journal of Mathematical and Statistical Psychology, 69*(3), 316-343. doi:10.1111/bmsp.12073

Rights, J., & Sterba, S. (2023). On the common but problematic specification of conflated random slopes in multilevel models. *Multivariate Behavioral Research, 58*(6), 1-28. doi:https://doi.org/10.1080/00273171.2023.2174490

Rights, J., Preacher, K., & Cole, D. (2020). The danger of conflating level-specific effects of control variables when primary interest lies in level-2 effects. *British Journal of Mathematical and Statistical Psychology, 73*, 194-211. doi:https://doi.org/10.1111/bmsp.12194

Shaw, M., & Flake, J. (2024). *Introduction to Multilevel Modelling.*

Snijders, T. (2005). Power and Sample Size in Multilevel Linear Models. In B. Everitt, Howell, & D (Eds.), *Encyclopedia of Statistics in Behavioral Science* (pp. 1570–1573). Wiley.

Snijders, T., & Bosker, R. (2012). *Multilevel Anlaysis. An Introduction to Basic and Advanced Multilevel Modeling.* Bodmin: SAGE.