Introduction to Two-Level Models for Clustered* Data

- Today's Class:
  - Fixed vs. Random Effects for Modeling Clustered Data
  - ICC and Design Effects in Clustered Data
  - Grand-Mean-Centering vs. Group-Mean Centering
  - Model Extensions under Group-MC and Grand-MC

** Clustering = Nesting = Grouping…*

Multilevel Models for Clustered Data

- So far we've seen multivariate models to model variance and covariance (dependency) arising from repeated measures (multiple conditions or trials from the same person)

- Now we examine multivariate (multilevel) models for more general examples of nesting/clustering/grouping:
  - Students within teachers, athletes within teams
  - Siblings within families, partners within dyads
  - Employees within businesses, patients within doctors

- Residuals of people from same group are likely to be correlated due to group differences (e.g., purposeful grouping or shared experiences create dependency)
  - Requires random effects → multiple piles of variance to predict

What is a Multilevel Model (MLM)?

- Same as other terms you have heard of:
  - **General Linear Mixed Model** (if you are from statistics)
    - *Mixed* = Fixed and Random effects
  - **Random Coefficients Model** (also if you are from statistics)
    - Random coefficients = Random effects
  - **Hierarchical Linear Model** (if you are from education)
    - Not the same as hierarchical regression

- Special cases of MLM:
  - Random Effects ANOVA or Repeated Measures ANOVA
  - (Latent) Growth Curve Model (where "Latent" → SEM)
  - Within-Person Variation Model (e.g., for daily diary data)
  - **Clustered/Nested Observations Model** (e.g., for kids in schools)
  - Cross-Classified Models (e.g., "value-added" models)

2 Options for Differences Across Groups

Represent Group Differences as Fixed Effects

- Include (#groups-1) contrasts for group membership in the **model for the means** (categorical X)→ so group is NOT another "level"
- Permits inference about differences between specific groups, but you cannot include between-group predictors (group is saturated)
- Snijders & Bosker (1999) ch. 4, p. 44 recommend if #groups < 10ish

Represent Group Differences as a Random Effect

- Include a **group random intercept variance in the model for the variance**, such that group differences become another "level"
- Permits inference about differences across groups more generally, for which you can test effects of between-group predictors
- Better if #groups > 10ish and you want to **predict** group differences

Empty Means, Random Intercept Model

MLM for Clustered Data:

- Two-level notation:
  - i = level 1, j = level 2
- Level 1:

$$y_{ij} = \beta_{0j} + e_{ij}$$

- Level 2:

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

Fixed Intercept =grand mean (of group means)

Random Intercept = group-specific deviation from predicted intercept

Residual = person-specific deviation from group's predicted outcome

3 Total Parameters:

Model for the Means (1):

- Fixed Intercept γ_{00}

Model for the Variance (2):

- Level-1 Variance of e_{ij} → σ_e^2
- Level-2 Variance of U_{0j} → $\tau_{U_0}^2$

Composite equation:

$$y_{ij} = (\gamma_{00} + U_{0j}) + e_{ij}$$

Intraclass Correlation (ICC)

Intraclass Correlation (ICC):

$$\text{ICC} = \frac{\text{BG}}{\text{BG} + \text{WG}} = \frac{\text{Intercept Variance}}{\text{Intercept Variance} + \text{Residual Variance}}$$

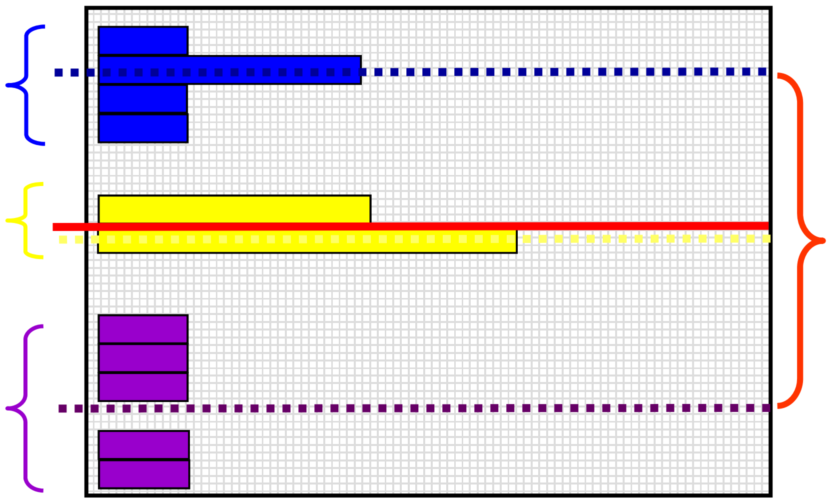
$$= \frac{\boldsymbol{\tau^2_{U_0}}}{\boldsymbol{\tau^2_{U_0}} + \boldsymbol{\sigma^2_e}}$$

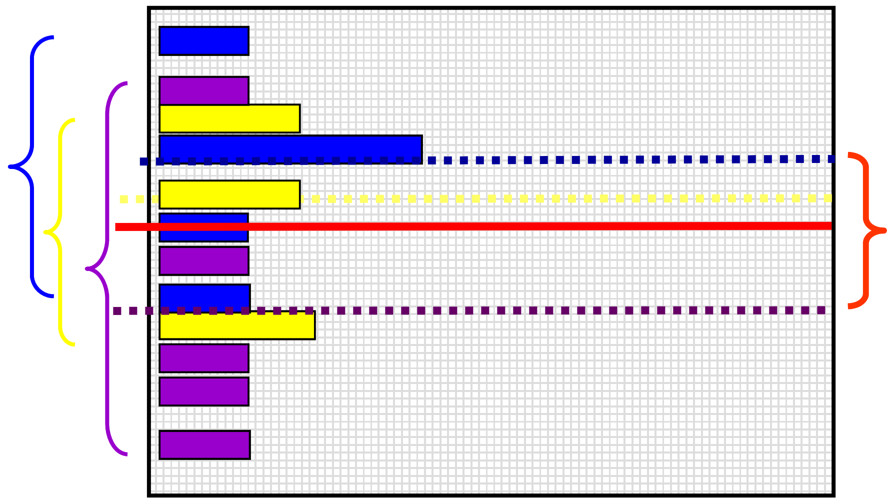
$\boldsymbol{\tau^2_{U_0}}$ → Why don't all groups have the same mean?
$\boldsymbol{\sigma^2_e}$ → Why don't all people from the same group have the same outcome?

- ICC = Proportion of total variance that is **between groups**
- ICC = Average **correlation** of persons from same group
- ICC is a standardized way of expressing how much we need to worry about *dependency due to group mean differences* **(i.e., ICC is an effect size for *constant* group dependency)**
- Whether $\boldsymbol{\tau^2_{U_0}}$>0 (whether ICC>0) is testable via a likelihood ratio test (-2ΔLL) against a general linear model with $\boldsymbol{\sigma^2_e}$ only

$$ICC = \frac{BetweenGroup}{BetweenGroup + WithinGroup} = \frac{\tau_{U_0}^2}{\tau_{U_0}^2 + \sigma_e^2}$$

Counter-Intuitive: Between-Group Variance is in the numerator, but the ICC is the correlation within a group!





ICC = BTW / BTW + within

→ Large ICC

→ Large within-group correlation

ICC = btw / btw + WITHIN

→ Small ICC

→ Small within-group correlation

Effects of Clustering on Effective N

- **Design Effect** expresses how much effective sample size needs to be adjusted due to clustering/grouping
- **Design Effect** = ratio of the variance using a given sampling design to the variance using a simple random sample from the same population, given the same total sample size either way

n = # persons from each group

- Design Effect = $1 + [(n-1) * \text{ICC}]$
- Effective sample size → $\text{N}_{\text{effective}} = \dfrac{\text{\# Total Observations}}{\text{Design Effect}}$

- As ICC goes UP and cluster size goes UP, the effective sample size goes DOWN
  - See Snijders and Bosker (2012) for more info and for a modified formula that takes unequal group sizes into account

Design Effects in Two-Level Nesting

- Design Effect = $1 + \left[(n-1) * \text{ICC}\right]$
- Effective sample size → $\text{N}_{\text{effective}} = \dfrac{\#\text{ Total Observations}}{\text{Design Effect}}$

- n=5 patients from each of 100 doctors, ICC=.30?
  - Patients Design Effect = 1 + (4 * .30) = 2.20
  - $\mathbf{N_{effective}}$ = 500 / 2.20 = **227** (not 500)

- n=20 students from each of 50 schools, ICC=.05?
  - Students Design Effect = 1 + (19 * .05) = 1.95
  - $\mathbf{N_{effective}}$ = 1000 / 1.95 = **513** (not 1000)

Does a non-significant ICC mean you can ignore groups and just do a regression?

- Effective sample size depends on BOTH the ICC and the number of people per group: As ICC goes UP and group size goes UP, the effective sample size goes DOWN
  - So there is NO VALUE OF ICC that is “safe” to ignore, not even ~0!
  - An ICC=0 in an *empty (unconditional)* model can become ICC>0 after adding person predictors because reducing the residual variance leads to an increase in the random intercept variance (→ *conditional* ICC > 0)

- So just do a multilevel analysis anyway…
  - Even if “that’s not your question”… because people come from groups, you still have to model group dependency appropriately because of:
    - Effect of clustering on fixed effect SE’s → **biased SEs**
    - Potential for **contextual effects** of person-level (level-1) predictors

Predictors in MLM for Clustered Data
Example: Achievement in Students nested in Schools

- <u>Level-2</u> predictors are <u>Group-Level</u> Variables
  - Can only have fixed effects (everyone gets the same effect)
  - e.g., Does mean school achievement differ b/t rural and urban schools?

- <u>Level-1</u> predictors are <u>Person-Level</u> Variables
  - Can have fixed, systematically varying, or random effects over groups
  - e.g., Does student achievement differ between boys and girls?
    - <u>Fixed effect:</u> Is there a gender difference in achievement, period?
    - <u>Systematically varying effect:</u> Does the gender effect differ b/t rural and urban schools? (but the gender effect is otherwise the same within rural and within urban schools)
    - <u>Random effect:</u> Does the gender effect differ *randomly* across schools? (even after including group-level predictors for it via cross-level interactions)

Level-1 (Person-Level) Predictors

- Modeling of level-1 predictors is complicated (and usually done incorrectly) because **each level-1 predictor is usually really 2 predictor variables** (each with their own effect), **not 1**

- Example: Student SES when students are clustered in schools
  - Some kids have more money than other kids in their school:
    - **WG variation in SES** *(represented directly as deviation from school mean)*
  - Some schools have more money than other schools:
    - **BG variation in SES** *(represented as school mean SES or via external info)*

- Can quantify each source of variance with an ICC
  - ICC = (BG variance) / (BG variance + WG variance)
  - ICC > 0? Level-1 predictor has BG variation (so it *could* have BG effect)
  - ICC < 1? Level-1 predictor has WG variation (so it *could* have WG effect)

Between-Group vs. Within-Group Effects

- Between-group and within-group effects in <u>SAME</u> direction
  - SES→ Achievement?
    - **BG: <u>Schools</u> with more money <u>than other schools</u> may have <u>greater</u> mean achievement than schools with less money**
    - **WG: <u>Kids</u> with more money <u>than other kids in their school</u> may have <u>greater</u> achievement than other kids in their school (regardless of school mean SES)**

- Between-group and within-group effects in <u>OPPOSITE</u> directions
  - Body mass → life expectancy in animals (Curran and Bauer, 2011)?
    - **BG: <u>Larger species</u> tend to have longer life expectancies than <u>smaller species</u> (e.g., whales live longer than cows, cows live longer than ducks)**
    - **WG: Within a species, <u>relatively bigger</u> animals have <u>shorter</u> life expectancy (e.g., fat ducks die sooner than skinny ducks)**

- Variables have **different meanings** and **different scales** across levels (so "one-unit" effects will rarely be the same across levels)!

Model Predictors (level-1 $x = x_{ij}$)

- **Level-2 effect of x_{ij}:**
  - The level-2 effect of x_{ij} can be represented by the group's mean across its persons of level-1 x_{ij} (labeled as $\mathbf{GMx_j}$ or $\bar{\mathbf{X}}_\mathbf{j}$)
  - $\mathbf{GMx_j}$ should be centered at a CONSTANT (grand mean or conceptually other useful value) so that 0 is meaningful, just like any other predictor

- **Level-1 effect of x_{ij} has two options:**
  - "**Grand-mean-centering (Grand-MC)**" → $\mathbf{L1x_{ij}} = \mathbf{x_{ij}} - \boldsymbol{C}$
    → level-1 predictor is centered using a CONSTANT
    (often but not necessarily the grand mean; it's just called that)
  - "**Group-mean-centering (Group-MC)**" → $\mathbf{WGx_{ij}} = \mathbf{x_{ij}} - \bar{\mathbf{X}}_\mathbf{j}$
    → level-1 predictor is centered using a level-2 VARIABLE
  - The interpretation of the level-2 effect of x_{ij} WILL DIFFER based on which centering method you choose for the level-1 effect of x_{ij}!
    - Grand-MC is more common in clustered data, so we'll use this

3 Kinds of Effects for Level-1 Predictors

- **Is the Within-Group (WG; level-1) fixed effect significant?**
  - If you have higher predictor values <u>than others in your group</u>, do you also have higher outcomes values <u>than others in your group</u>, such that the within-group deviation $\mathbf{WGx_{ij}}$ accounts for level-1 residual variance (σ_e^2)?
  - Given directly by the level-1 effect of WGx_{ij} if using Group-MC —*OR* — given directly by the level-1 effect of $L1x_{ij}$ if using Grand-MC and including GMx_j at level 2 (without GMx_j, the level-1 effect of $L1x_{ij}$ if using Grand-MC is the smushed effect)

- **Is the Between-Group (BG; level-2) fixed effect significant?**
  - Are groups with higher predictor values <u>than other groups</u> also higher on Y <u>than other groups</u>, such that the group mean of the person-level predictor $\mathbf{GMx_j}$ accounts for level-2 random intercept variance ($\tau_{U_0}^2$)?
  - Given directly by level-2 effect of GMx_j if using Group-MC for the level-1 predictor (or can be requested via ESTIMATE if using Grand-MC for the level-1 predictor)

- **Are the BG and WG effects different: Is there a level-2 contextual effect?**
  - After controlling for the absolute value of the level-1 predictor for each person, is there still <u>an incremental contribution from the group mean of the predictor</u> (i.e., does a group's general tendency predict $\tau_{U_0}^2$ above and beyond the person-specific predictor value)?
  - Given directly by level-2 effect of GMx_j if using Grand-MC for the level-1 predictor (or can be requested via ESTIMATE if using Group-MC for the level-1 predictor)

Contextual Effects in Clustered Data

- Grand-MC can be more convenient in clustered data due to its ability to directly provide level-2 contextual effects (that are often of interest)
- Example: Effect of SES for students (nested in schools) on achievement:

- **Grand-MC** of level-1 student $\mathrm{SES_{ij}}$, school mean $\overline{\mathrm{SES}}_\mathrm{j}$ included at level 2
  - At level-1 → $\mathbf{L1\_SES_{ij}} = \mathrm{SES_{ij}} - C$ At level-2 → $\mathbf{BG\_SES_{ij}} = \overline{\mathrm{SES}}_\mathrm{j} - C$
  - Level-1 **WG** effect: Effect of being rich kid relative to your school (is purely WG after *statistically* controlling for $\overline{\mathrm{SES}}_\mathrm{j}$)
  - Level-2 **Contextual** effect: Incremental effect of going to a rich school (after *statistically* controlling for kid $\mathrm{SES_{ij}}$)

- **Group-MC** of level-1 student $\mathrm{SES_{ij}}$, school mean $\overline{\mathrm{SES}}_\mathrm{j}$ included at level 2
  - At level-1 → $\mathbf{WG\_SES_{ij}} = \mathrm{SES_{ij}} - \overline{\mathrm{SES}}_\mathrm{j}$ At level-2 → $\mathbf{BG\_SES_{ij}} = \overline{\mathrm{SES}}_\mathrm{j} - C$
  - Level-1 **WG** effect: Effect of being rich kid relative to your school (is already purely WG because of centering around $\overline{\mathrm{SES}}_\mathrm{j}$)
  - Level-2 **BG** effect: Effect of going to a rich school NOT controlling for kid $\mathrm{SES_{ij}}$

WRONG WAY: Clustered Data Model with **x_{ij} represented at Level 1 Only:**

→ WG and BG Effects are **Smushed Together**

x_{ij} is grand-mean-centered into $L1x_{ij}$, WITHOUT GMx_j at L2:

Level 1: $y_{ij} = \beta_{0j} + \beta_{1j}(\mathbf{L1x_{ij}}) + e_{ij}$

$L1x_{ij} = x_{ij} - C$ → it still has both Level-2 BG and Level-1 WG variation

Level 2: $\beta_{0j} = \gamma_{00} + U_{0j}$

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

γ_{10} = *smushed* WG and BG effects

Because $L1x_{ij}$ still contains its original 2 different kinds of variation (BG and WG), its 1 fixed effect has to do the work of 2 predictors!

A *smushed* effect is also referred to as the *convergence*, *conflated*, or *composite* effect

Convergence (Smushed) Effect of a Level-1 Predictor

$$\text{Convergence Effect: } \gamma_{\text{conv}} \approx \frac{\dfrac{\gamma_{BG}}{SE^2_{BG}} + \dfrac{\gamma_{WG}}{SE^2_{WG}}}{\dfrac{1}{SE^2_{BG}} + \dfrac{1}{SE^2_{WG}}}$$

Adapted from Raudenbush & Bryk (2002, p. 138)

- **The convergence effect will often be closer to the within-group effect** (due to larger level-1 sample size and thus smaller SE)

- **It is the rule, not the exception, that between and within effects differ** (Snijders & Bosker, 1999, p. 52-56, and personal experience!)

- However—when grand-mean-centering a level-1 predictor, **convergence is testable** by including a **contextual effect (carried by the group mean)** for how the **BG effect** differs from the **WG effect**…

Clustered Data Model with **Grand-Mean-Centered Level-1 x_{ij}**

→ Model tests difference of WG vs. BG effects (It's been fixed!)

x_{ij} is grand-mean-centered into $L1x_{ij}$, <u>WITH</u> GMx_j at L2:

Level 1: $y_{ij} = \beta_{0j} + \beta_{1j}(L1x_{ij}) + e_{ij}$

$L1x_{ij} = x_{ij} - C$ → it still has both Level-2 BG and Level-1 WG variation

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + U_{0j}$

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

$GMx_j = \bar{X}_j - C$ → it has only Level-2 BG variation

γ_{10} becomes the WG effect → *unique* level-1 effect after controlling for GMx_j

γ_{01} *becomes* the contextual effect that indicates <u>how the BG effect differs from the WG effect</u>
→ *unique* level-2 effect after controlling for $L1x_{ij}$
→ does group matter beyond individuals?

Clustered Data Model with **Group-Mean-Centered Level-1 x_{ij}**

→ WG and BG Effects directly through **separate** parameters

x_{ij} is group-mean-centered into WGx_{ij}, with GMx_j at L2:

Level 1: $y_{ij} = \beta_{0j} + \beta_{1j}(WGx_{ij}) + e_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + U_{0j}$

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

$WGx_{ij} = x_{ij} - \bar{X}_j$ → it has only Level-1 WG variation

$GMx_j = \bar{X}_j - C$ → it has only Level-2 BG variation

γ_{10} = WG main effect of having more x_{ij} than others in your group

γ_{01} = BG main effect of having more $\bar{X}_j$ than other groups

Because WGx_{ij} and GMx_j are uncorrelated, each gets the total effect for its level (WG=L1, BG=L2)

Why the Difference? Remember Regular Old Regression?

- In this model: $Y_i = \beta_0 + \beta_1(X_{1i}) + \beta_2(X_{2i}) + e_i$
- If X_{1i} and X_{2i} **ARE NOT** correlated:
  - β_1 is **ALL the relationship** between X_{1i} and Y_i
  - β_2 is **ALL the relationship** between X_{2i} and Y_i

- If X_{1i} and X_{2i} **ARE** correlated:
  - β_1 is **different than** the full relationship between X_{1i} and Y_i
    - "Unique" effect of X_{1i} *controlling for* X_{2i} or *holding* X_{2i} *constant*
  - β_2 is **different than** the full relationship between X_{2i} and Y_i
    - "Unique" effect of X_{2i} *controlling for* X_{1i} or *holding* X_{1i} *constant*

- Hang onto that idea…

Why the Difference? Group-MC vs. Grand-MC Variable for Level-1 Predictors

Level 2		Original	Group-MC Level 1	Grand-MC Level 1
$\bar{X}_j$	$GMx_j = \bar{X}_j - 5$	x_{ij}	$WGx_{ij} = x_{ij} - \bar{X}_j$	$L1x_{ij} = x_{ij} - 5$
3	−2	2	−1	−3
3	−2	4	1	−1
7	2	6	−1	1
7	2	8	1	3

Same GMx_j goes into the model using either way of centering the level-1 variable x_{ij}

Using **Group-MC**, WGx_{ij} has NO level-2 BG variation, so it is not correlated with GMx_j

Using **Grand-MC**, $L1x_{ij}$ STILL has level-2 BG variation, so it is STILL CORRELATED with GMx_j

So the effects of GMx_j and $L1x_{ij}$ when included together under Grand-MC will be different than their effects would be if they were by themselves…

Group-MC and Grand-MC Models are Equivalent Given a Fixed Level-1 Main Effect Only

Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij} - GMx_j) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + U_{0j}$

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

→ $y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + e_{ij}$

→ $y_{ij} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$

Grand-MC: $L1x_{ij} = x_{ij}$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij}) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + U_{0j}$

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

→ $y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$

Composite Model:
← As Group-MC
← As Grand-MC

Effect	Group-MC	Grand-MC
Intercept	γ_{00}	γ_{00}
WG Effect	γ_{10}	γ_{10}
Contextual	$\gamma_{01} - \gamma_{10}$	γ_{01}
BG Effect	γ_{01}	$\gamma_{01} + \gamma_{10}$

Variance Accounted For By Level-2 Predictors

- **Fixed effects of level 2 predictors *by themselves*:**
  - Level-2 (BG) main effects reduce level-2 (BG) random intercept variance
  - Level-2 (BG) interactions also reduce level-2 (BG) random intercept variance

- **Fixed effects of *cross-level interactions* (level 1* level 2):**
  - If the interacting level-1 predictor is <u>random</u>, any cross-level interaction with it will reduce its corresponding level-2 BG random slope variance (that line's U)
  - If the interacting level-1 predictor <u>not random</u>, any cross-level interaction with it will reduce the level-1 WG residual variance instead
    - This is because the level-2 BG random slope variance would have been created by decomposing the level-1 residual variance in the first place
    - The level-1 effect would then be called "**systematically varying**" to reflect a compromise between "fixed" (all the same) and "random" (all different)—it's not that each group needs their own slope, but that the slope varies systematically across groups as a function of a known group predictor (and not otherwise)

Variance Accounted For By Level-1 Predictors

- **Fixed effects of level 1 predictors *by themselves*:**
  - Level-1 (WG) main effects reduce Level-1 (WG) residual variance
  - Level-1 (WG) interactions also reduce Level-1 (WG) residual variance

- **What happens at level 2 depends on what kind of variance the level-1 predictor has:**
  - If the level-1 predictor ALSO has level-2 variance (e.g., Grand-MC predictors), then its level-2 variance will also likely reduce level-2 random intercept variance
  - If the level-1 predictor DOES NOT have level-2 variance (e.g., Group-MC predictors), then its reduction in the level-1 residual variance will cause an INCREASE in level-2 random intercept variance
    - Same thing happens with Grand-MC level-1 predictors, but you don't generally see it
  - It's just an artifact that the estimate of true random intercept variance is:

    True $\tau^2_{U_0}$ = observed $\tau^2_{U_0} - \frac{\sigma^2_e}{n}$ → so if only σ^2_e decreases, $\tau^2_{U_0}$ increases

The Joy of Interactions Involving Level-1 Predictors

- <u>Must consider interactions with both its BG and WG parts:</u>
- Example: Does the effect of employee motivation (x_{ij}) on employee performance interact with type of business (for profit or non-profit; $Type_j$)?

- <u>Group-Mean-Centering:</u>
  - $\mathbf{WGx_{ij}} * \mathbf{Type_j}$ → Does the WG motivation effect differ between business types?
  - $\mathbf{GMx_j} * \mathbf{Type_j}$ → Does the BG motivation effect differ between business types?
    - Moderation of <u>total</u> group motivation effect (not controlling for individual motivation)
    - If forgotten, then $\mathbf{Type_j}$ moderates the motivation effect only at level 1 (WG, not BG)

- <u>Grand-Mean-Centering:</u>
  - $\mathbf{L1x_{ij}} * \mathbf{Type_j}$ → Does the WG motivation effect differ between business types?
  - $\mathbf{GMx_j} * \mathbf{Type_j}$ → Does the *contextual* motivation effect differ b/t business types?
    - Moderation of <u>incremental</u> group motivation effect <u>controlling for employee motivation</u> (moderation of the "boost" in group performance from working with motivated people)
    - If forgotten, then although the level-1 main effect of motivation has been un-smushed via the main effect of $\mathbf{GMx_j}$, the interaction of $\mathbf{L1x_{ij}} * \mathbf{Type_j}$ would still be smushed

Interactions with Level-1 Predictors:
Example: Employee Motivation (x_{ij}) by Business Type ($Type_j$)

Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij} - GMx_j) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + U_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11}(Sex_j)$

Composite: $y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + e_{ij}$
$+ \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + \gamma_{11}(Type_j)(x_{ij} - GMx_j)$

Grand-MC: $L1x_{ij} = x_{ij}$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij}) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + U_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11}(Type_j)$

Composite: $y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$
$+ \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + \gamma_{11}(Type_j)(x_{ij})$

Interactions Involving Level-1 Predictors Belong at Both Levels of the Model

On the left below → Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

$y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + e_{ij}$
$+ \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + \gamma_{11}(Type_j)(x_{ij} - GMx_j)$

← **As Group-MC**

$y_{ij} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$
$+ \gamma_{02}(Type_j) + (\gamma_{03} - \gamma_{11})(Type_j)(GMx_j) + \gamma_{11}(Type_j)(x_{ij})$

← **As Grand-MC**

On the right below → Grand-MC: $L1x_{ij} = x_{ij}$

$y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$
$+ \gamma_{02}(Type_j) + \gamma_{03}(Type_j)(GMx_j) + \gamma_{11}(Type_j)(x_{ij})$

After adding an interaction for $Type_j$ with x_{ij} at both levels, then the Group-MC and Grand-MC models are equivalent

---

Intercept: $\gamma_{00} = \gamma_{00}$ **BG Effect:** $\gamma_{01} = \gamma_{01} + \gamma_{10}$ **Contextual:** $\gamma_{01} = \gamma_{01} - \gamma_{10}$

WG Effect: $\gamma_{10} = \gamma_{10}$ **BG*Type Effect:** $\gamma_{03} = \gamma_{03} + \gamma_{11}$ **Contextual*Type:** $\gamma_{03} = \gamma_{03} - \gamma_{11}$

Type Effect: $\gamma_{20} = \gamma_{20}$ **BG*WG or Contextual*WG is the same:** $\gamma_{11} = \gamma_{11}$

Intra-variable Interactions

- Still must consider interactions with both its BG and WG parts!
- Example: Does the effect of employee motivation (x_{ij}) on employee performance interact with business group mean motivation (GMx_j)?

- Group-Mean-Centering:
  - $\mathbf{WGx_{ij}} * \mathbf{GMx_j}$ → Does the WG motivation effect differ by group motivation?
  - $\mathbf{GMx_j} * \mathbf{GMx_j}$ → Does the BG motivation effect differ by group motivation?
    - Moderation of total group motivation effect (not controlling for individual motivation)
    - If forgotten, then $\mathbf{GMx_j}$ moderates the motivation effect only at level 1 (WG, not BG)

- Grand-Mean-Centering:
  - $\mathbf{L1x_{ij}} * \mathbf{GMx_j}$ → Does the WG motivation effect differ by group motivation?
  - $\mathbf{GMx_j} * \mathbf{GMx_j}$ → Does the *contextual* motivation effect differ by group motiv.?
    - Moderation of incremental group motivation effect controlling for employee motivation (moderation of the boost in group performance from working with motivated people)
    - If forgotten, then although the level-1 main effect of motivation has been un-smushed via the main effect of $\mathbf{GMx_j}$, the interaction of $\mathbf{L1x_{ij}} * \mathbf{GMx_j}$ would still be smushed

Intra-variable Interactions:

Example: Employee Motivation (x_{ij}) by Business Mean Motivation (GMx_j)

Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij} - GMx_j) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{02}(GMx_j)(GMx_j) + U_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11}(GMx_j)$

Composite: $y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + e_{ij}$
$+ \gamma_{02}(GMx_j)(GMx_j) + \gamma_{11}(GMx_j)(x_{ij} - GMx_j)$

Grand-MC: $L1x_{ij} = x_{ij}$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij}) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{02}(GMx_j)(GMx_j) + U_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11}(GMx_j)$

Composite: $y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij}$
$+ \gamma_{02}(GMx_j)(GMx_j) + \gamma_{11}(GMx_j)(x_{ij})$

Intra-variable Interactions:

Example: Employee Motivation (x_{ij}) by Business Mean Motivation (GMx_j)

On the left below → Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

$$y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + e_{ij} + \gamma_{02}(GMx_j)(GMx_j) + \gamma_{11}(GMx_j)(x_{ij} - GMx_j)$$

← As Group-MC

$$y_{ij} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij} + (\gamma_{02} - \gamma_{11})(GMx_j)(GMx_j) + \gamma_{11}(GMx_j)(x_{ij})$$

← As Grand-MC

On the right below → Grand-MC: $L1x_{ij} = x_{ij}$

$$y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + e_{ij} + \gamma_{02}(GMx_j)(GMx_j) + \gamma_{11}(GMx_j)(x_{ij})$$

After adding an interaction for $Type_j$ with x_{ij} at both levels, then the Group-MC and Grand-MC models are equivalent

Intercept: $\gamma_{00} = \gamma_{00}$ **BG Effect:** $\gamma_{01} = \gamma_{01} + \gamma_{10}$ **Contextual:** $\gamma_{01} = \gamma_{01} - \gamma_{10}$

WG Effect: $\gamma_{10} = \gamma_{10}$ **BG² Effect:** $\gamma_{02} = \gamma_{02} + \gamma_{11}$ **Contextual²:** $\gamma_{02} = \gamma_{02} - \gamma_{11}$

BG*WG or Contextual*WG is the same: $\gamma_{11} = \gamma_{11}$

When Group-MC ≠ Grand-MC: Random Effects of Level-1 Predictors

Group-MC: $WGx_{ij} = x_{ij} - GMx_j$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij} - GMx_j) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + U_{0j}$

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

Variance due to GMx_j is removed from the random slope in Group-MC.

$\rightarrow y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij} - GMx_j) + U_{0j} + U_{1j}(x_{ij} - GMx_j) + e_{ij}$

Grand-MC: $L1x_{ij} = x_{ij}$

Level-1: $y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij}) + e_{ij}$

Level-2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(GMx_j) + U_{0j}$

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

Variance due to GMx_j is still part of the random slope in Grand-MC. So these models cannot be made equivalent.

$\rightarrow y_{ij} = \gamma_{00} + \gamma_{01}(GMx_j) + \gamma_{10}(x_{ij}) + U_{0j} + U_{1j}(x_{ij}) + e_{ij}$

Random Effects of Level-1 Predictors

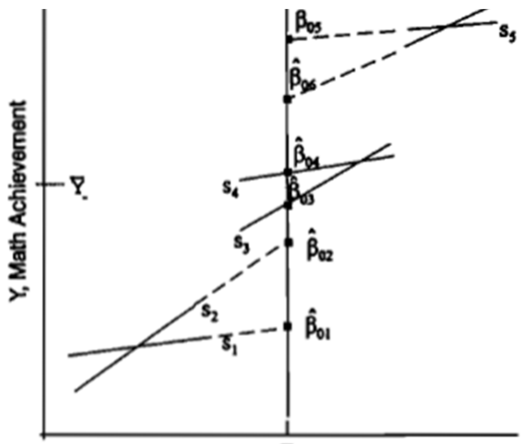
- **Random intercepts** mean different things under each model:
  - **Group-MC** → Group differences at **WGx_{ij} =0** (that every group has)
  - **Grand-MC** → Group differences at **$L1x_{ij}$=0** (that not every group will have)

- **Differential shrinkage of the random intercepts** results from differential reliability of the intercept data across models:
  - Group-MC → Won't affect shrinkage of slopes unless highly correlated
  - Grand-MC → Will affect shrinkage of slopes due to forced extrapolation

- As a result, the **random slope variance may be smaller** under Grand-MC than under Group-MC
  - Problem worsens with greater ICC of level-1 predictor (more extrapolation)
  - Anecdotal example was presented in Raudenbush & Bryk (2002; chapter 5)

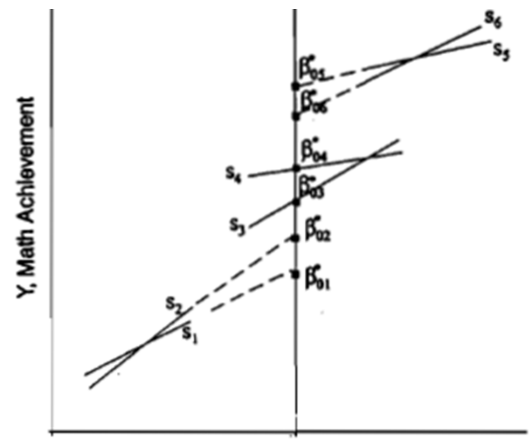
Bias in Random Slope Variance

OLS Per-Group Estimates



Level-1 X

EB Shrunken Estimates



Level-1 X

Top right: Intercepts and slopes are homogenized in Grand-MC because of intercept extrapolation

Bottom: Downwardly-biased random slope variance in Grand-MC relative to Group-MC

Unconditional Results	*Conditional Results*
Group-MC	
$\hat{\mathbf{T}} = \begin{bmatrix} 8.68 & 0.05 \\ 0.05 & 0.68 \end{bmatrix}$ $\hat{\sigma}^2 = 36.70$	$\hat{\mathbf{T}} = \begin{bmatrix} 2.38 & 0.19 \\ 0.19 & 0.15 \end{bmatrix}$ $\hat{\sigma}^2 = 36.70$
Grand-MC	
$\hat{\mathbf{T}} = \begin{bmatrix} 4.83 & -0.15 \\ -0.15 & 0.42 \end{bmatrix}$ $\hat{\sigma}^2 = 36.83$	$\hat{\mathbf{T}} = \begin{bmatrix} 2.41 & 0.19 \\ 0.19 & 0.06 \end{bmatrix}$ $\hat{\sigma}^2 = 36.74$

MLM for Clustered Data: Summary

- Models now come in only two kinds: “empty” and “conditional”
  - The lack of a comparable dimension to “time” simplifies things greatly!
- L2 = Between-Group, L1 = Within-Group (between-person)
  - Level-2 predictors are group variables: can have fixed or systematically varying effects (but not random effects in two-level models)
  - Level-1 predictors are person variables: can have fixed, random, or systematically varying effects
- No smushing main effects or interactions of level-1 predictors:
  - Group-MC at Level 1: Get L1=WG and L2=BG effects directly
  - Grand-MC at Level 1: Get L1=WG and L2=contextual effects directly
    - As long as some representation of the L1 effect is included in L2; otherwise, the L1 effect (and any interactions thereof) will be smushed