Three-Level Random Effects Models for Longitudinal Data

- Topics:
  - Example three-level designs for occasions, persons, clusters
  - Partitioning variation across three levels in clustered longitudinal data (occasions within persons within clusters)
    - Bonus: Also for three-level longitudinal (occasions, days, persons)
  - Unconditional (time only) models for clustered longitudinal
    - Bonus: Also for time-varying groups instead
  - Conditional (+ predictors) models for clustered longitudinal
    - Variable-centering vs constant-centering; pseudo-R^2

2 Options for Differences Across Clusters

Represent Cluster Dependency as Fixed Effects

- Include (#clusters−1) binary predictors for cluster membership in the **model for the means** → **so cluster is NOT a model "level"**
  - Main effects control for cluster mean differences only; interactions with person predictors are also needed to control for cluster slope differences
- Useful if #clusters < 10ish or you care about specific clusters, but then you cannot include cluster predictors → saturated mean diffs

Represent Cluster Differences via Random Effects

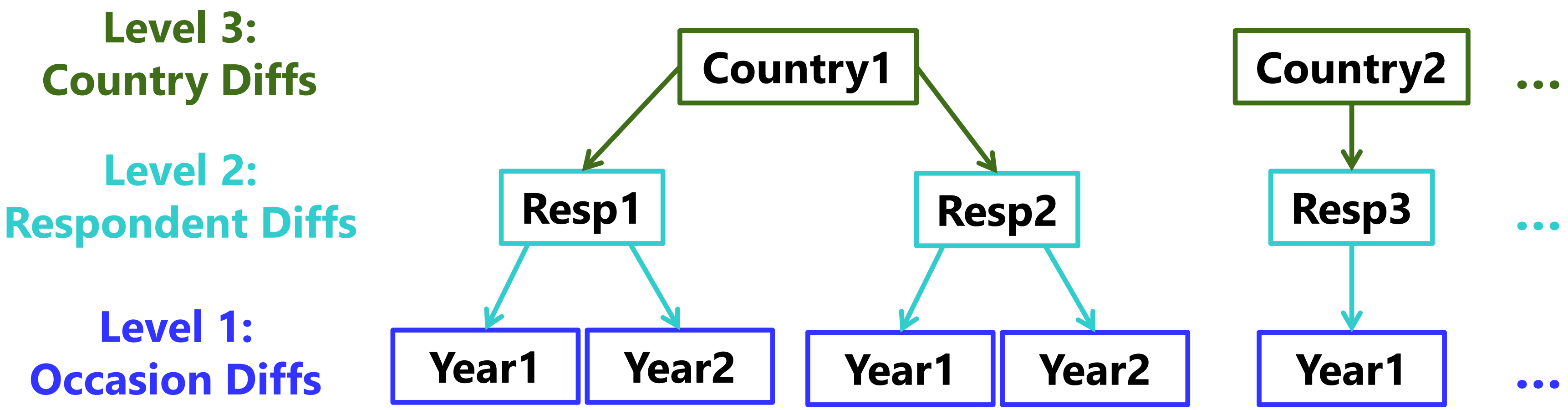
- Include a random intercept variance across clusters in the **model for the variance** → **then cluster IS a new model "level"**
  - A random intercept controls for cluster mean differences only; a random slope variance is needed for cluster differences in person predictor slopes
- Better if #clusters > 10ish or you want to **predict** cluster differences

What determines the number of levels?

- **Answer: the model for the outcome variance ONLY**
- How many dimensions of sampling in the outcome?
  - Longitudinal, one person per family? → 2-level model
  - Longitudinal, 2+ people per family? → 3-level model
  - Longitudinal, 2+ people per family, many cities? → 4-level model
  - Sampling dimensions may also be crossed instead of nested, or may be modeled with fixed effects if the # units is small
- Need at least one pile of variance per dimension (for 3 levels, that's 2 levels of random effects and a residual)
  - Include whatever predictors you want per level, but keep in mind that the usefulness of your predictors will be constrained by the amount of outcome variance in its relevant sampling dimension

Kinds of 3-Level Designs: Clustered Longitudinal

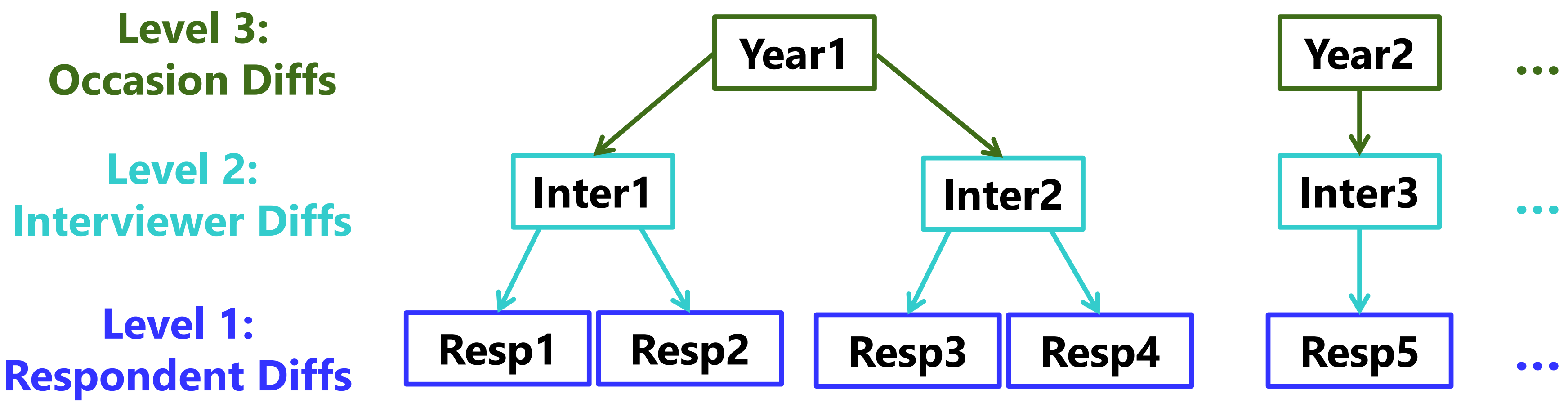
- First example: Predicting **occasion-specific respondent outcomes** for people nested in countries, collected over several years (**all same people** and **same countries** are measured over time)



- Country predictors can be included at level 3 only (no random effects)
- Person predictors should be included at levels 2 and 3 (+random over 3)
- What about effects of **time-varying predictors**?
  - For **People**: effects should be included at all 3 levels (+random over 2 and 3)
  - For **Countries**: effects are only possible at levels 1 and 3 (+random over 3)

Other Examples of 3-Level Designs

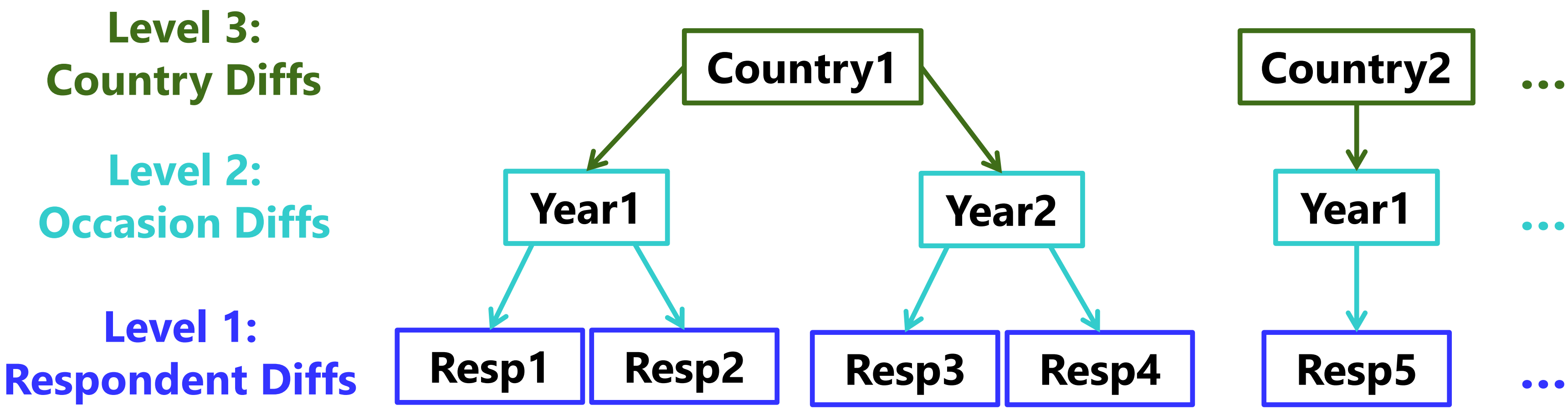
- The sampling design for the outcome (not the predictors) dictates what your levels will be, **so occasion may not always be level 1**
- Example: Predicting answer compliance in respondents nested in interviewers, collected over several years (**all different people**)



- Based on # occasions, occasion mean differences may be modeled…
  - As fixed effects in the model for the means → 2-level model instead
    - Best to use dummy codes for time if few occasions OR no time-level predictors of interest
  - Via a random intercept in the model for the variance → 3-level model
    - Then differences in compliance rates over time can be predicted by time-level predictors

Other Examples of 3-Level Designs

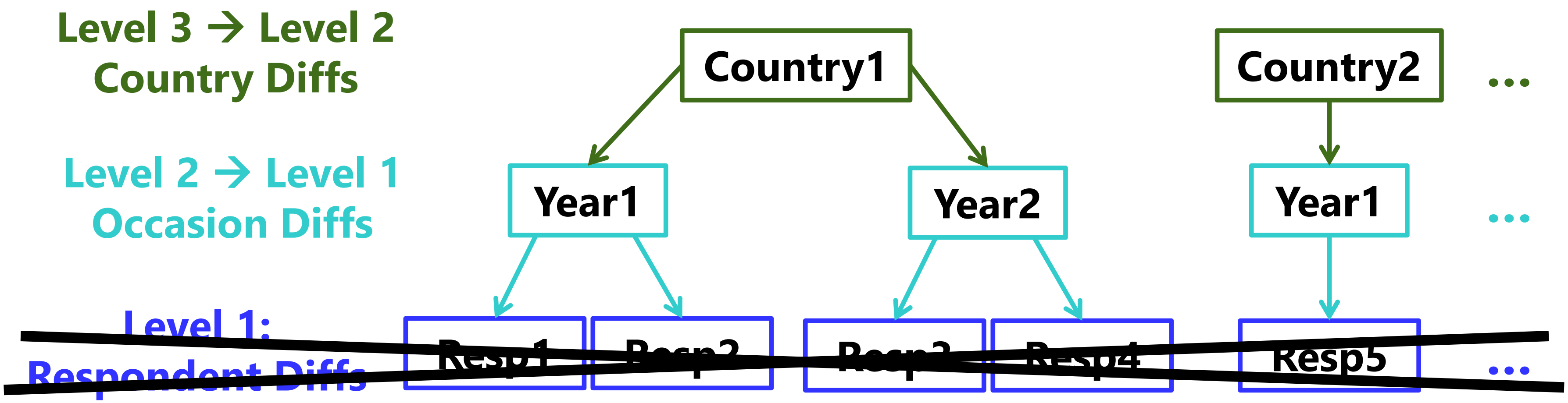
- Another example: Predicting **occasion-specific respondent outcomes** for people nested in countries, collected over several years (**all different people**, but the **same countries** measured over time)



- Before including any fixed effects of time, the dimensions of country and time are **actually crossed**, not nested as shown here
  - Are nested **after** controlling for occasion mean differences via fixed effects (using dummy codes per mean or a time trend that describes the means)
  - Time is still level 2 because not all countries change/fluctuate the same way

3-Level Designs: Predictors vs. Outcomes

- Same example: What if, instead of respondent outcomes, we wanted to predict **time-varying country outcomes**?



If the outcome is measured at level 2 (per country per time):

- Respondents are no longer a level at all (no outcomes for them)
- So there is nothing for respondent predictors to do, except at higher levels
  - **Time-specific averages** of respondent predictors → time-level outcome variation
  - **Across time, country averages** of respondent predictors → country-level outcome variation

Empty Means, 3-Level Random Intercept Model: Example for Clustered Longitudinal Data

Notation: t = level-1 time, i = level-2 person, c = level-3 cluster

Level 1: $y_{tic} = \beta_{0ic} + e_{tic}$

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

Level 2: $\beta_{0ic} = \delta_{00c} + U_{0ic}$

Person Random Intercept = person-specific deviation from cluster's predicted outcome

Level 3: $\delta_{00c} = \gamma_{000} + V_{00c}$

Fixed Intercept =grand mean of cluster means

Cluster Random Intercept = cluster-specific deviation from fixed intercept

4 Total Parameters:

Model for the Means (1):

- Fixed Intercept γ_{00}

Model for the Variance (2):

- Level-1 Variance of e_{tic} → σ^2_e
- Level-2 Variance of U_{0ic} → $\tau^2_{U_0}$
- Level-3 Variance of V_{00c} → $\tau^2_{V_{00}}$

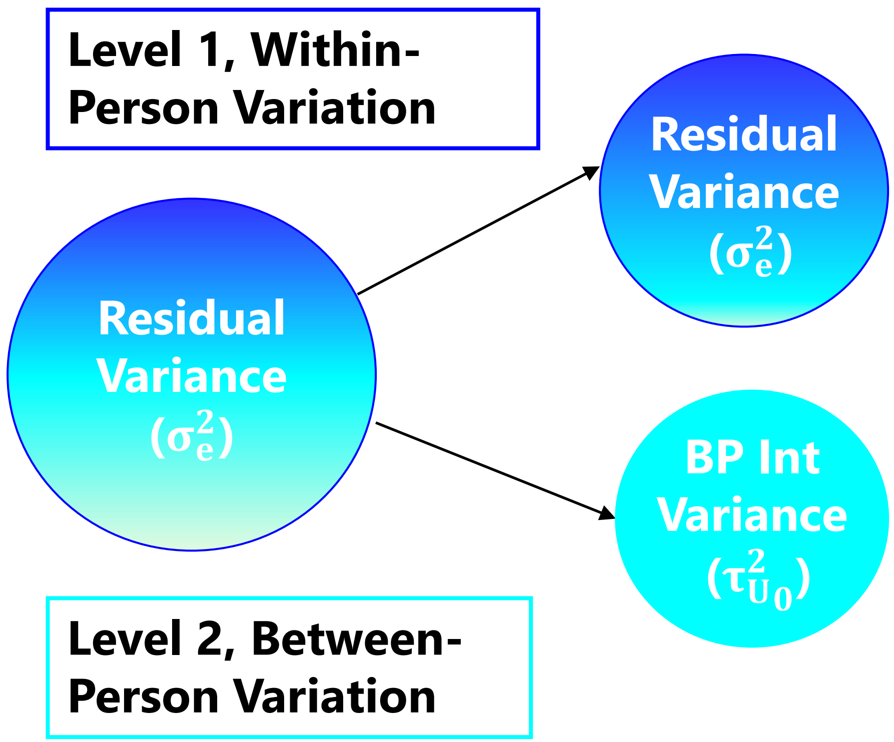
Composite equation:

$$y_{tic} = \gamma_{000} + V_{00c} + U_{0ic} + e_{tic}$$

Btw: My bad for reusing "V"

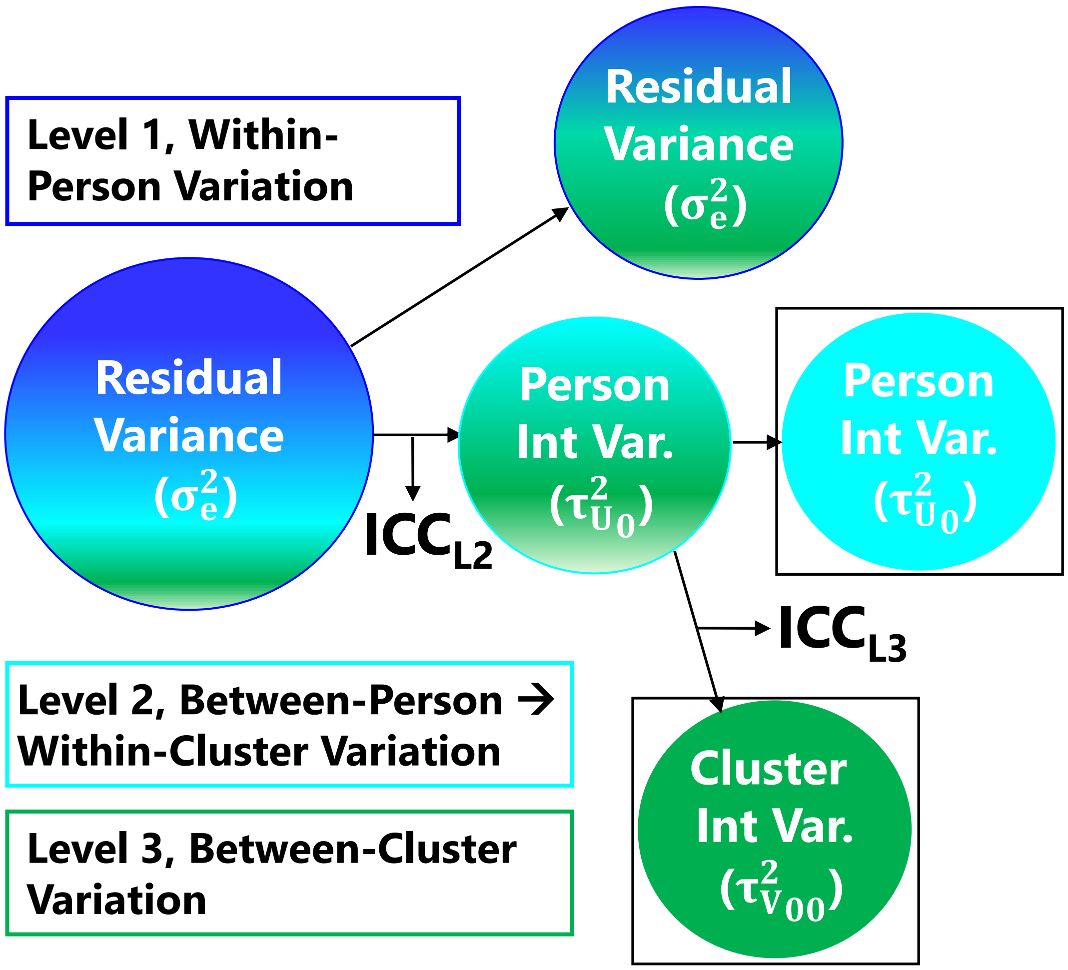
Example 2-Level Random Intercept Model

- Where does each kind of person dependency go? Into a new random effects variance component (or "pile" of variance):
- Let's start with an empty means, random intercept 2-level model for occasions within persons:



Example 3-Level Random Intercept Model

- Now let's move to an empty means, random intercept 3-level model of occasions within persons within clusters:



ICCs in a 3-Level Random Intercept Model: Occasions within Persons within Clusters

- **ICC for level 2 (and level 3) relative to level 1:**

- $$\text{ICC}_{\text{L2}} = \frac{\text{Between−Person}}{\text{Total}} = \frac{\text{L3+L2}}{\text{L3+L2+L1}} = \frac{\boldsymbol{\tau}^2_{\mathbf{V_{00}}} + \boldsymbol{\tau}^2_{\mathbf{U_0}}}{\boldsymbol{\tau}^2_{\mathbf{V_{00}}} + \boldsymbol{\tau}^2_{\mathbf{U_0}} + \boldsymbol{\sigma}^2_{\mathbf{e}}}$$

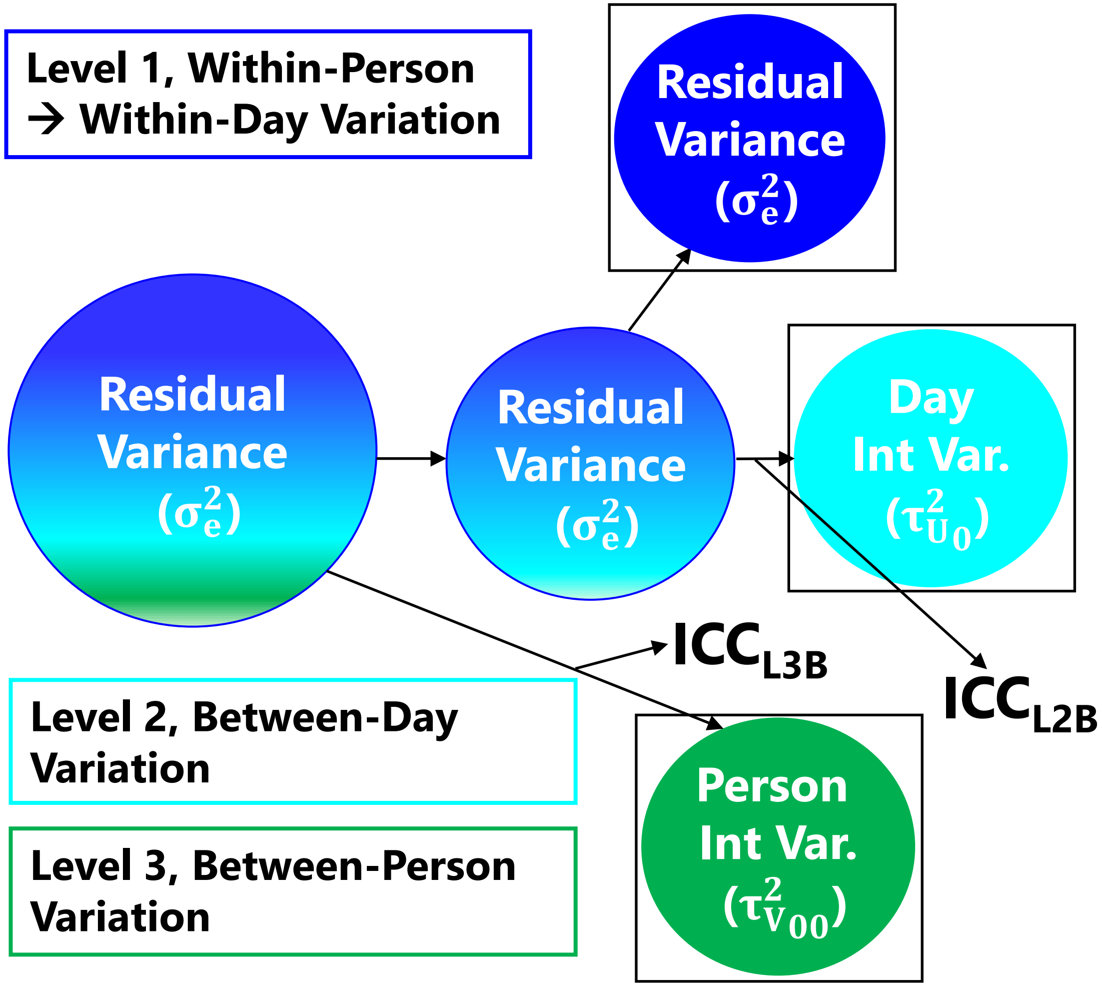
  → This ICC expresses the similarity of **occasions from the same person** (and by definition, from the same cluster) → of the ***total outcome variation***, how much of it is ***between persons, or cross-sectional (not due to time)***?

- **ICC for level 3 relative to level 2 (ignoring level 1):**

- $$\text{ICC}_{\text{L3}} = \frac{\text{Between−Cluster}}{\text{Between−Person}} = \frac{\text{L3}}{\text{L3+L2}} = \frac{\boldsymbol{\tau}^2_{\mathbf{V_{00}}}}{\boldsymbol{\tau}^2_{\mathbf{V_{00}}} + \boldsymbol{\tau}^2_{\mathbf{U_0}}}$$

  → This ICC expresses the similarity of **persons from the same cluster** (ignoring within-person variation over time) → of ***that total between-person outcome variation***, how much of that is actually ***between clusters***?

Bonus: 3-Level Model for Intensive Longitudinal Data (occasions, days, persons)



Useful ICC variants for this type of design:

ICC_{L3B} = L3 / total

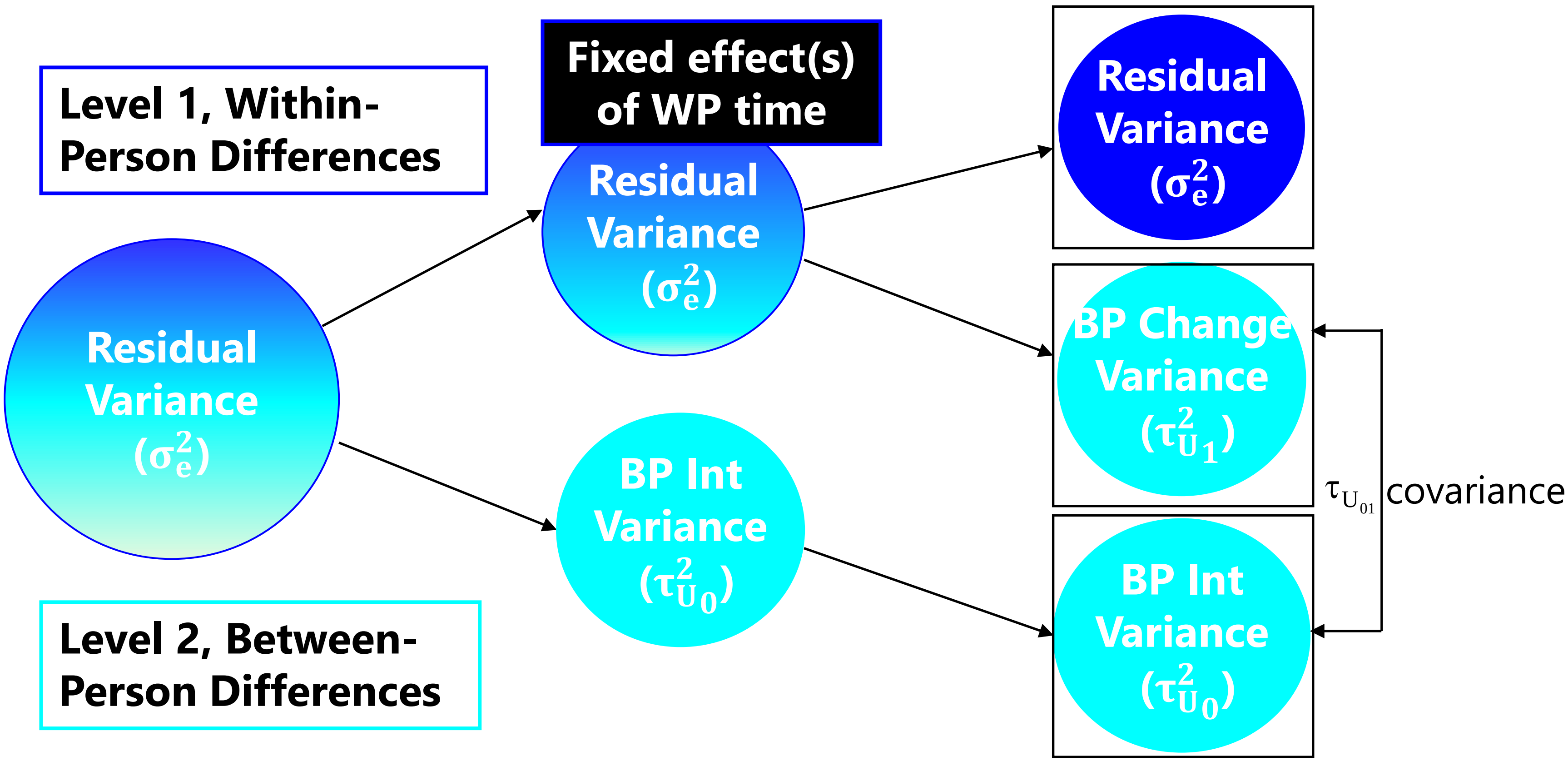
- % Between Persons
- Note: this is what is given by STATA and Mplus as "level-3 ICC"

ICC_{L2B} = L2 / (L2+L1)

- Proportion of time-related variance for day
- Tests if occasions on same day are more related than occasions on different days (i.e., is day needed?)

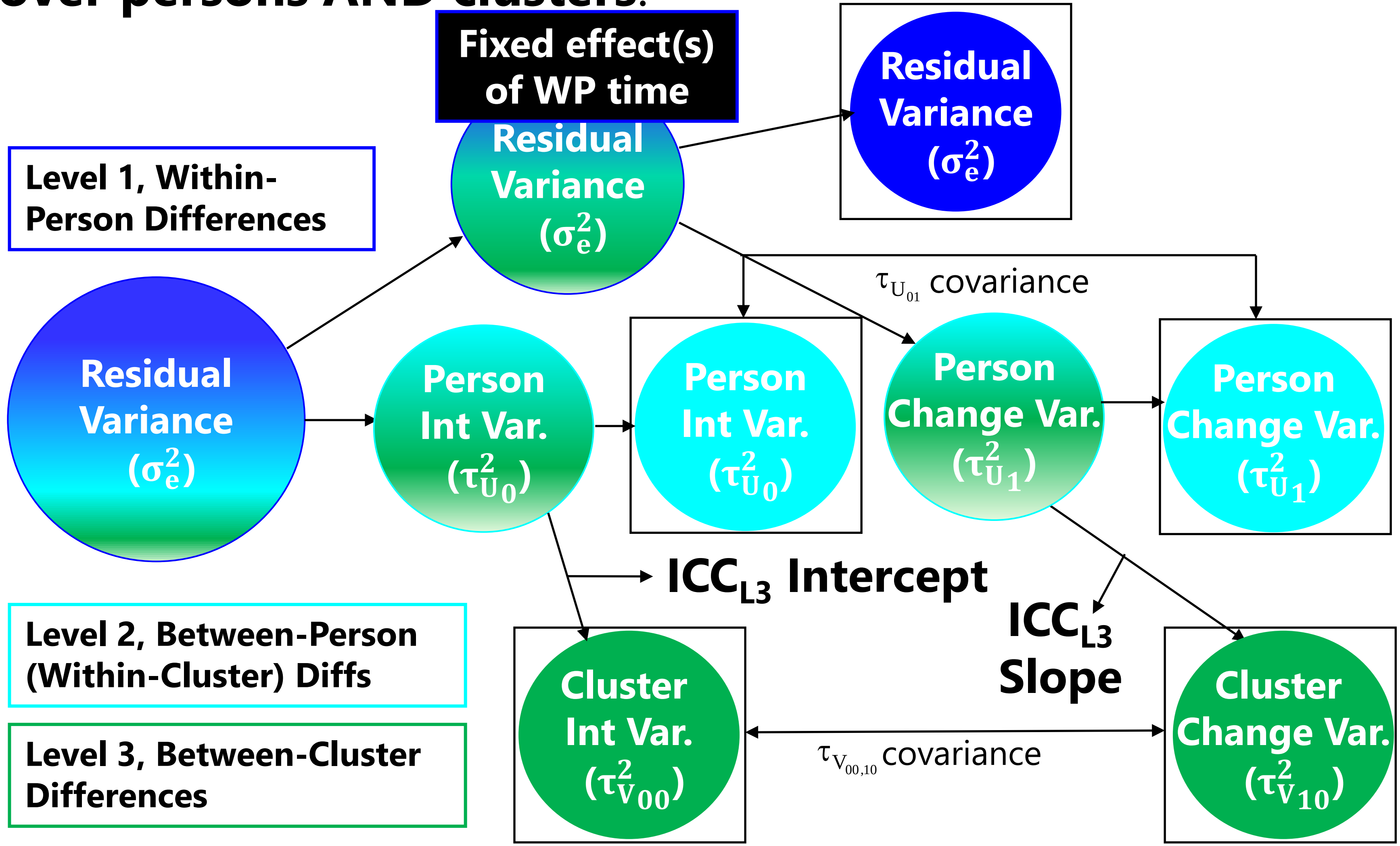
Example 2-Level Random Change Model

- What about time? After adding fixed effects of time, we can add random effects of time over persons in a 2-level model:



Example 3-Level Random Change Model

- In a 3-level model, we can have **random effects of time over persons AND clusters**:



Example 3-Level Random Change Model

Notation: t = level-1 time, i = level-2 person, c = level-3 cluster

Level 1: $y_{tic} = \beta_{0ic} + \beta_{1ic}(Time_{tic}) + e_{tic}$

Residual = time-specific deviation from **person's** predicted growth line (σ_e^2)

Level 2: $\beta_{0ic} = \delta_{00c} + U_{0ic}$

$\beta_{1ic} = \delta_{10c} + U_{1ic}$

Person Random Intercept and Slope = person-specific deviations from **cluster's** predicted intercept, change ($\tau_{U_0}^2$, $\tau_{U_1}^2$, $\tau_{U_{01}}$)

Level 3: $\delta_{00c} = \gamma_{000} + V_{00c}$

$\delta_{10c} = \gamma_{100} + V_{10c}$

Cluster Random Intercept and Change = cluster-specific deviations from fixed intercept, change ($\tau_{V_{00}}^2$, $\tau_{V_{10}}^2$, $\tau_{V_{00,10}}$)

Fixed Intercept, Fixed Linear Change Slope

Composite equation (9 parameters):

$$y_{tic} = (\gamma_{000} + V_{00c} + U_{0ic}) + (\gamma_{100} + V_{10c} + U_{1ic})(Time_{tic}) + e_{tic}$$

Random Time Slopes at both Levels 2 AND 3? An example with family as cluster:

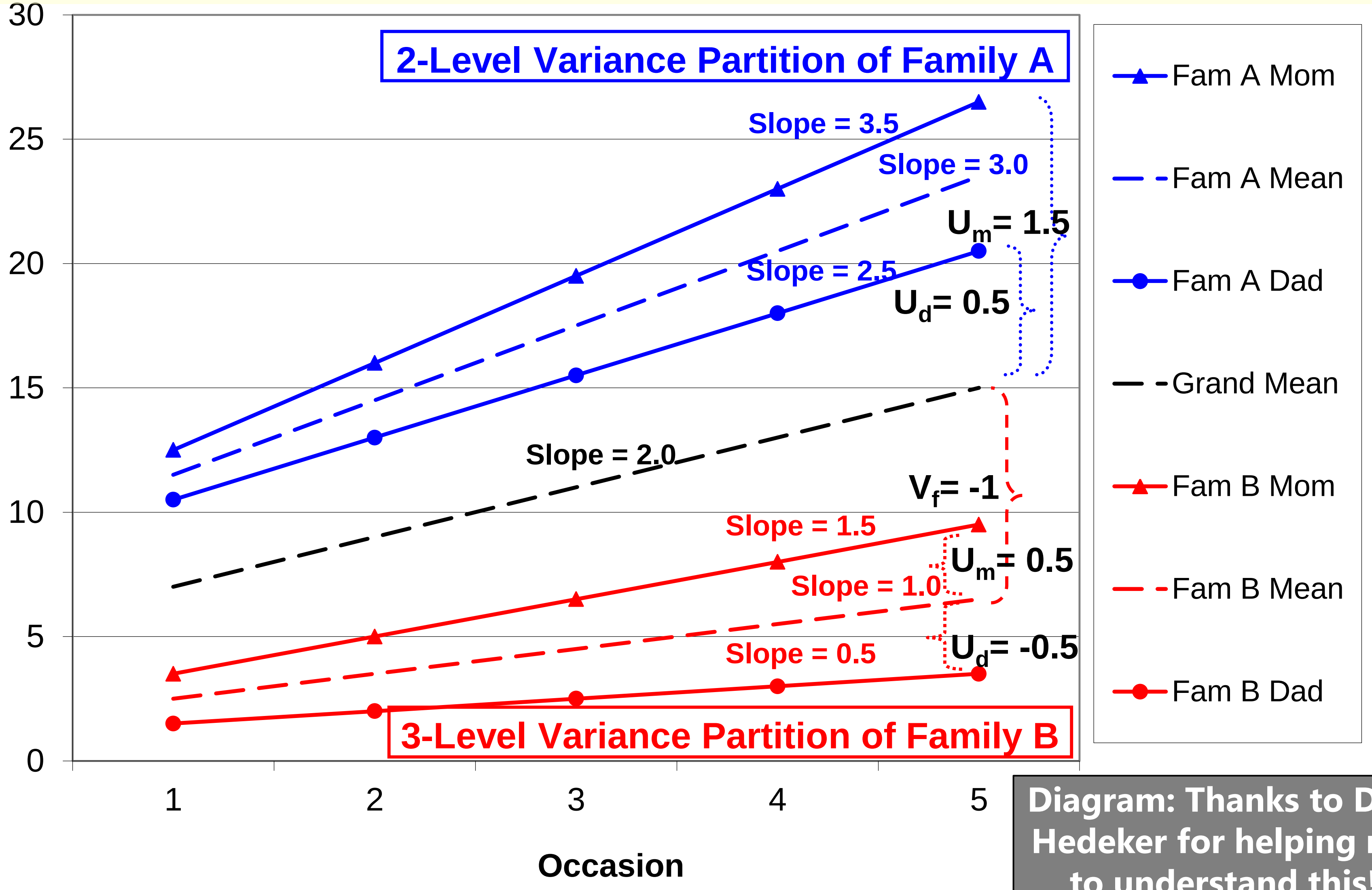


Diagram: Thanks to Don Hedeker for helping me to understand this!

ICCs for Random Intercepts and Slopes

- Once random time slopes are included at both level-3 and level-2, ICCs can be computed for the random intercepts and time slopes specifically (which is the level-3 type of ICC)

$$\text{ICC}_{\text{Int}} = \frac{\text{Between} - \text{Cluster}}{\text{Between} - \text{Person}} = \frac{\text{L3 Int}}{\text{L3 Int} + \text{L2 Int}} = \frac{\boldsymbol{\tau}^2_{\mathbf{V_{00}}}}{\boldsymbol{\tau}^2_{\mathbf{V_{00}}} + \boldsymbol{\tau}^2_{\mathbf{U_0}}}$$

$$\text{ICC}_{\text{Slope}} = \frac{\text{Between} - \text{Cluster}}{\text{Between} - \text{Person}} = \frac{\text{L3 Slope}}{\text{L3 Slope} + \text{L2 Slope}} = \frac{\boldsymbol{\tau}^2_{\mathbf{V_{10}}}}{\boldsymbol{\tau}^2_{\mathbf{V_{10}}} + \boldsymbol{\tau}^2_{\mathbf{U_1}}}$$

- Can be computed for any level-1 slope that is random at both levels (e.g., linear and quadratic time, time-varying predictors)
- Be careful when the model is uneven across levels, though

$$\frac{\text{Random Level 2: int, linear, quad}}{\text{Random Level 3: int, linear}} \rightarrow \frac{\text{Linear is when } \mathbf{time} = \mathbf{0}}{\text{Linear is at } \mathbf{any} \text{ occasion}}$$

More on Random Slopes in 3-Level Models

- Any level-1 predictor can have a random slope over level 2, level 3, or over both levels at once, but I recommend working your way **UP the higher levels** for assessing random effects…
  - e.g., Does the effect of time vary over level-2 persons?
  - If so, does the effect of time vary over level-3 clusters, too? → Is there a **commonality** in how people from the same cluster change over time?
- … because random effects at level 3 only are possible but unlikely (e.g., means everyone in the cluster changes the same)

- Level-2 predictors can also have random effects over level 3
  - e.g., Does the effect of a L2 person characteristic vary over L3 clusters?
- Level-1, level-2, and level-1 by level-2 cross-level interactions can all have random effects over level 3, too, in theory
  - But tread carefully! The more random effects you have, the more likely you are to have convergence problems ("**G** matrix not positive definite")

What about Time-Varying Clusters?

- e.g., Students are nested within classes at each occasion…
- But if students move into different classes over time…
  - Level-1 occasions are nested within level-2 students AND within level-2 classes: Students are crossed with classes at level 2
- How to model a **time-varying classroom effect**?
  - Btw, this is the basis of so-called "value-added models"

- Two example options (both via fixed or random effects):
  - **"Acute" effect**: Class effect active only when students are in that class
    - e.g., class effect ← teacher bias
    - Once a student is out of the class, class effect is no longer present
  - **"Transfer" effect**: Effect is active when in class AND in the future…
    - e.g., class effect ← differential learning
    - Effect stays with the student in the future (i.e., a "layered" value-added model)

Time (t), Students (s), and Classes (c)

- Custom-built intercepts for time-varying effects of classes
  - An intercept is usually a column of 1's, but ours will be 0's and 1's to serve as switches that turn on/off class effects

				Per-Year Class ID (−99 = missing)			Intercepts for Acute Effects			Intercepts for Transfer Effects		
Student ID	Class ID	Grade	Year	Year 0 Class	Year 1 Class	Year 2 Class	Year 0 Intercept	Year 1 Intercept	Year 2 Intercept	Year 0 Effect	Year 1 Effect	Year 2 Effect
101	1	3	0	1	−99	43	1	0	0	1	0	0
101	−99	4	1	1	−99	43	0	0	0	0	0	0
101	43	5	2	1	−99	43	0	0	1	1	0	1
102	3	3	0	3	21	42	1	0	0	1	0	0
102	21	4	1	3	21	42	0	1	0	1	1	0
102	42	5	2	3	21	42	0	0	1	1	1	1

Time (t), Students (s), and Classes (c)

- Hoffman (2015) Equation 11.3: **fixed effects model** for classroom as a categorical time-varying predictor:
  - Allows for control of classroom differences only….

$$
\begin{aligned}
\text{Effort}_{tsc} &= \gamma_{000} + \gamma_{100}\left(\text{Year01}_{tsc}\right) + \gamma_{200}\left(\text{Year12}_{tsc}\right) + U_{0s0} + e_{tsc} \\
&+ \gamma^{0}_{001}\left(\text{Class1}_{c}\right)\left(\text{Int0}_{tsc}\right) + \gamma^{0}_{002}\left(\text{Class2}_{c}\right)\left(\text{Int0}_{tsc}\right) \cdots + \gamma^{0}_{00C}\left(\text{ClassC}_{c}\right)\left(\text{Int0}_{tsc}\right) \\
&+ \gamma^{1}_{001}\left(\text{Class1}_{c}\right)\left(\text{Int1}_{tsc}\right) + \gamma^{1}_{002}\left(\text{Class2}_{c}\right)\left(\text{Int1}_{tsc}\right) \cdots + \gamma^{1}_{00C}\left(\text{ClassC}_{c}\right)\left(\text{Int1}_{tsc}\right) \\
&+ \gamma^{2}_{001}\left(\text{Class1}_{c}\right)\left(\text{Int2}_{tsc}\right) + \gamma^{2}_{002}\left(\text{Class2}_{c}\right)\left(\text{Int2}_{tsc}\right) \cdots + \gamma^{2}_{00C}\left(\text{ClassC}_{c}\right)\left(\text{Int2}_{tsc}\right)
\end{aligned}
$$

- Hoffman (2015) Equation 11.4: classrooms as a random effect crossed with students (as a random effect) at level 2:
  - Controls and quantifies classroom variance so it can be predicted!

$$
\begin{aligned}
\text{Effort}_{tsc} &= \gamma_{000} + \gamma_{100}\left(\text{Year01}_{tsc}\right) + \gamma_{200}\left(\text{Year12}_{tsc}\right) + U_{0s0} + e_{tsc} \\
&+ U^{0}_{00c}\left(\text{Int0}_{tsc}\right) + U^{1}_{00c}\left(\text{Int1}_{tsc}\right) + U^{2}_{00c}\left(\text{Int2}_{tsc}\right)
\end{aligned}
$$

Clustered Longitudinal Data: Conditional Model Specification

- Remember separating between- and within-person effects?
  Now there are **three** potential fixed effects for any **level-1** predictor!
  - Example in a Clustered Longitudinal Design: Effect of stress on wellbeing, both measured over time within person within families:
  - **Level 1** (Time): During **Occasions** of more stress, people have lower (time-specific) wellbeing than during occasions of less stress
  - **Level 2** (Person): **People** in the family who have more stress have lower (person average) wellbeing than people in the family who have less stress
  - **Level 3** (Family): **Families** who have more stress have lower (family average) wellbeing than families who have less stress

- And **two** potential fixed effects for any **level-2** predictor:
  - Example: Effect of baseline level of person coping skills in same design:
  - **Level 2** (Person): **People** in the family who cope better have better (person average) wellbeing than people in the family who cope worse
  - **Level 3** (Family): **Families** who cope better have better (family average) wellbeing than families who cope worse

3 Options to Prevent Smushed Slopes

- Within Univariate MLM framework (predict only one column):
  1. **Variable-centering**: Carve up lower-level predictor into its level-specific parts using observed variables (1 predictor per *relevant* level)
     - Will always yield **per-level total slopes** (that explain variance at its level only
  2. **Constant-centering**: Do NOT carve up lower-level predictor, but add relevant upper-level means to distinguish each upper-level slope
     - Choice of constant is still irrelevant (changes where 0 is, not what variance it has)
     - Will always yield **lowest-level within slope** and **upper-level contextual slopes**!
     - Do NOT do this if you want a **random lowest-level slope** (→ random smushed)
- Within Multivariate MLM framework (via M-SEM or SEM):
  3. **Latent-centering**: Lower-level predictor is another outcome → let the model carve it up into **level-specific latent variables**
     - Best in theory, but the type of upper-level slope provided (between or contextual) depends on type of model syntax (and the estimator in Mplus)! (Hoffman, 2019)
     - **Multivariate MLM is the only recommended option for level-1 predictors that contain individual differences in change over time at any upper level!**

Option 1: Separate Total Effects Per Level Using **Variable-Centering**

- **Level 1 (Occasion):** *Time-varying stress relative to person mean*
  - → $WPstress_{tic}$ = $Stress_{tic}$ – $PersonMeanStress_{ic}$
  - → Directly tests if within-person effect ≠ 0?
  - → **Total** within-person effect of more stress ***than usual*** ≠ 0?

Only ok if TV stress does not have random diffs in change over time

- **Level 2 (Person):** *Person mean stress relative to family (within family)*
  - → $WFstress_{ic}$ = $PersonMeanStress_{ic}$ – $FamilyMeanStress_{c}$
  - → Directly tests if within-family effect ≠ 0?
  - → **Total** effect of more stress ***than other members of one's family*** ≠ 0?

- **Level 3 (Family):** *Family mean stress relative to all families (from constant)*
  - → $BFstress_{c}$ = $FamilyMeanStress_{c}$ – C
  - → Directly tests if between-family effect ≠ 0?
  - → **Total** effect of more stress ***than other families*** ≠ 0?

Option 1: Separate Total Effects Per Level Using **Variable-Centering**

Notation: t = level-1 time, i = level-2 person, c = level-3 family
PM = person mean, FM = family mean, C = centering constant

Level 1: $y_{tic} = \beta_{0ic} + \beta_{1ic}(Time_{tic}) + \beta_{2ic}(Stress_{tic} - PMstress_{ic}) + e_{tic}$

Level 2: $\beta_{0ic} = \delta_{00c} + \delta_{01c}(PMstress_{ic} - FMstress_{c}) + U_{0ic}$

$\beta_{1ic} = \delta_{10c} + U_{1ic}$

$\beta_{2ic} = \delta_{20c} + (U_{2ic})$

Level 3: $\delta_{00c} = \gamma_{000} + \gamma_{001}(FMstress_{c} - C_{3}) + V_{00c}$ — **Fixed intercept, Between-family stress main effect**

$\delta_{01c} = \gamma_{010} + (V_{01c})$ — **Within-family stress main effect**

$\delta_{10c} = \gamma_{100} + V_{10c}$ — **Time main effect**

$\delta_{20c} = \gamma_{200} + (V_{20c})$ — **Within-person stress main effect**

Option 2: Contextual Effects Per Level Using **Constant-Centering**

- **Level 1 (Occasion):** *Time-varying stress (relative to sample constant)*
  - → TVstress_{tic} = $\text{Stress}_{tic} - C_1$
  - → Directly tests if within-person effect ≠ 0?
  - → **Total** within-person effect of more stress ***than usual*** ≠ 0?

Only ok if TV stress does not have random diffs in change over time and it will have fixed slopes only

- **Level 2 (Person):** *Person mean stress (relative to sample constant)*
  - → BPstress_{ic} = $\text{PersonMeanStress}_{ic} - C_2$
  - → Directly tests if within-person and within-family effects ≠ ?
  - → **Contextual** effect of more stress ***than other members of one's family*** ≠ 0?

- **Level 3 (Family):** *Family mean stress relative to all families (from constant)*
  - → BFstress_{c} = $\text{FamilyMeanStress}_{c} - C_3$
  - → Directly tests if within-family and between-family effects ≠ ?
  - → **Contextual** effect of more stress ***than other families*** ≠ 0?

Option 2: Contextual Effects Per Level Using **Constant-Centering**

Notation: t = level-1 time, i = level-2 person, c = level-3 family
PM = person mean, FM = family mean, C = centering constant

Level 1: $y_{tic} = \beta_{0ic} + \beta_{1ic}(Time_{tic}) + \beta_{2ic}(Stress_{tic} - C_1) + e_{tic}$

Level 2: $\beta_{0ic} = \delta_{00c} + \delta_{01c}(PMstress_{ic} - C_2) + U_{0ic}$

$\beta_{1ic} = \delta_{10c} + U_{1ic}$

$\beta_{2ic} = \delta_{20c} + (U_{2ic})$

Level 3: $\delta_{00c} = \gamma_{000} + \gamma_{001}(FMstress_{c} - C_3) + V_{00c}$ — **Fixed intercept, <u>Contextual</u> family stress main effect**

$\delta_{01c} = \gamma_{010} + (V_{01c})$ — **<u>Contextual</u> within-family stress main effect**

$\delta_{10c} = \gamma_{100} + V_{10c}$ — **Time main effect**

$\delta_{20c} = \gamma_{200} + (V_{20c})$ — **Within-person stress main effect**

What does it mean to omit higher-level effects under each centering method?

- **Variable-Centering**: Omitting a fixed effect assumes that the effect at that level **does not exist** (= 0 like usual for no slope)
  - Remove L3 effect? Assume L3 Between-Family effect = 0
    - *L1 effect = Within-Person effect, L2 effect = Within-Family effect*
  - Then remove L2 effect? Assume L2 Within-Family effect = 0
    - *L1 effect = Within-Person effect*

- **Constant-Centering**: Omitting a fixed effect means the effect at that level **is equivalent to** the effect at **the level below**
  - Remove L3 effect? Assume L3 Between-Family = L2 Within-Family effect
    - *L1 effect = Within-Person effect, L2 effect = smushed WF and BF effects*
  - Then remove L2 effect? Assume L2 Between-Person effect = L1 effect
    - *L1 smushed = Within-Person, Within-Family, and Between-Family effects*

Btw, interactions belong at each level, too…

- Example: Is the effect of stress on wellbeing moderated by time-invariant person coping? Using **variable-centering**:
- **Stress Effects**

Adding random slopes? Should use variable-centering (or latent-centering)

  - **Level 1 (Occasion):** $WPstress_{tic} = Stress_{tic} - PersonMeanStress_{ic}$
  - **Level 2 (Person):** $WFstress_{ic} = PersonMeanStress_{ic} - FamilyMeanStress_{c}$
  - **Level 3 (Family):** $BFstress_{c} = FamilyMeanStress_{c} - C_3$
- **Coping Effects**
  - **Level 2 (Person):** $WFcope_{ic} = Cope_{ic} - FamilyMeanCope_{c}$
  - **Level 3 (Family):** $BFcope_{c} = FamilyMeanCope_{c} - C_3$
- **Interaction Effects**
- With level-1 stress: $WPstress_{tic}$ * $WFcope_{ic}$, $WPstress_{tic}$ * $BFcope_{c}$
- With level-2 stress: $WFstress_{ic}$ * $WFcope_{ic}$, ($WFstress_{ic}$ * $BFcope_{c}$)
- With level-3 stress: $BFstress_{c}$ * $BFcope_{c}$, ($BFstress_{c}$ * $WFcope_{ic}$)

Btw, interactions belong at each level, too…

Notation: t = level-1 time, i = level-2 person, c = level-3 family
PM = person mean, FM = family mean, C = centering constant

Level 1: $y_{tic} = \beta_{0ic} + \beta_{1ic}(Time_{tic}) + \beta_{2ic}(Stress_{tic} - PMstress_{ic}) + e_{tic}$

Level 2: $\beta_{0ic} = \delta_{00c} + \delta_{01c}(PMstress_{ic} - FMstress_c)$
$+ \delta_{02c}(Cope_{ic} - FMcope_c)$
$+ \delta_{03c}(PMstress_{ic} - FMstress_c)(Cope_{ic} - FMcope_c) + U_{0ic}$

$\beta_{1ic} = \delta_{10c} + U_{1ic}$

$\beta_{2ic} = \delta_{20c} + \delta_{21c}(Cope_{ic} - FMcope_c) + (U_{2ic})$

Level 3: $\delta_{00c} = \gamma_{000} + \gamma_{001}(FMstress_c - C_3) + \gamma_{002}(FMcope_c - C_3)$
$+ \gamma_{003}(FMstress_c - C_3)(FMcope_c - C_3) + V_{00c}$

$\delta_{01c} = \gamma_{010} + (V_{01c})$ $\delta_{02c} = \gamma_{020} + (V_{02c})$ $\delta_{03c} = \gamma_{030} + (V_{03c})$

$\delta_{10c} = \gamma_{100} + V_{10c}$

$\delta_{20c} = \gamma_{200} + \gamma_{202}(FMcope_c - C) + (V_{20c})$ $\delta_{21c} = \gamma_{210} + (V_{21c})$

Pseudo-R^2 in Three-Level Models

- Although it may not work this neatly in real data, here is the logic for how each type of fixed slope should explain variance
- **Main effects** and purely **same-level interactions** are straightforward—they target their **own level**:
  - L1 main effects and L1 interactions → L1 residual variance
  - L2 main effects and L2 interactions → L2 random intercept variance
  - L3 main effects and L3 interactions → L3 random intercept variance
- For **cross-level interactions**, which variance gets explained **depends** on if **random slopes** are included at each level…
  - L3 * L1 → L3 random variance in L1 slope if included, or L2 random variance in L1 slope if included, or L1 residual otherwise
  - L3 * L2 → L3 random variance in L2 slope if included, or L2 random intercept otherwise
  - L2 * L1 → L2 random variance in L1 slope if included, or L1 residual otherwise

Summary: Three-Level Random Effects Models

- Estimating 3-level models requires no new concepts, but everything is an order of complexity higher:
  - **Partitioning variance** over 3 levels instead of 2 → **many possible ICCs**
  - **Random slope variance will come from the variance directly below:**
    - Level-2 random slope variance comes from level-1 residual
    - Level-3 random slope variance comes from level-2 random slope (or residual)
  - **Level-1 effects can be random over level 2, level 3, or both at once**
    - ICCs can be computed for level-1 slopes that are random over both level-2 and level-3 (assuming the L2 and L3 variance models match)
    - Smushing of level-1 fixed effects should be tested over levels 2 AND 3
  - **Level-2 effects can be random over level 3**
    - Smushing of level-2 fixed effects should be tested over level 3
  - **Level-3 effects cannot be random**; no worries about smushing
  - **Pseudo-R^2** follows similar patterns as for two-level models
  - Phew….!