Time-Varying Predictors for Within-Person Fluctuation

• Today’s topics:
  - Review of time-invariant predictors
  - Time-varying predictors that fluctuate over time
  - Person-Mean-Centering (PMC)
  - Grand-Mean-Centering (GMC)
  - Model extensions under Person-MC vs. Grand-MC
  - Model assumptions
  - Predicting heterogeneity of variance
Modeling Time-Invariant Predictors

What independent variables can be time-invariant predictors?

• Also known as “person-level” or “level-2” predictors
• Include substantive predictors, controls, and predictors of missingness

• Can be anything that **does not change across time** (e.g., Biological Sex)

• Can be anything that **is not likely to change across the study**, but you may have to argue for this (e.g., Parenting Strategies, SES)

• Can be anything that **does change across the study**...
  • But you have **only measured once**
    • Limit conclusions to variable’s status at time of measurement
    • e.g., “Parenting Strategies at age 10”
  • Or **is perfectly correlated with time** (age, time to event)
    • Would use Age at Baseline, or Time to Event from Baseline instead
The Role of Time-Invariant Predictors in the Model for the Means

- In Within-Person Change Models → Adjust growth curve

Main effect of X, No interaction with time

Interaction with time, Main effect of X?

Main effect of X, and Interaction with time
The Role of Time-Invariant Predictors in the Model for the Means

- In **Within-Person Fluctuation Models** → Adjust mean level

![Graph showing no main effect of X and main effect of X with time](image)
Why Level-2 Predictors Cannot Have Random Effects in 2-Level Models

Random Slopes for Time

Time
(or Any Level-1 Predictor)

Random Slopes for Sex?

Sex
(or any Level-2 Predictor)

You cannot make a line out of a dot, so level-2 effects cannot vary randomly over persons.
Education as a Time-Invariant Predictor: Example using a Random Quadratic Time Model

• **Main Effect of Education = Education*Intercept Interaction**
  - Moderates the intercept → Increase or decrease in expected outcome at time 0 for every year of education

• **Effect of Education on Linear Time = Education*Time Interaction**
  - Moderates the linear time slope → Increase or decrease in expected rate of change at time 0 for every year of education

• **Effect of Education on Quadratic Time = Education*Time² Interaction**
  - Moderates the quadratic time slope → Increase or decrease in half of expected acceleration/deceleration of linear rate of change for every year of education
Education (12 years = 0) as a Time-Invariant Predictor: Example using a Random Quadratic Time Model

Level 1: \( y_{ti} = \beta_{0i} + \beta_{1i} \text{Time}_{ti} + \beta_{2i} \text{Time}_{ti}^2 + e_{ti} \)

Level 2 Equations (one per \( \beta \)):

\[ \beta_{0i} = \gamma_{00} + \gamma_{01} Ed_i + U_{0i} \]
- Intercept for person \( i \)
- Fixed Intercept when Time=0 and Ed=12
- \( \Delta \) in Intercept per unit \( \Delta \) in Ed
- Random (Deviation) Intercept after controlling for Ed

\[ \beta_{1i} = \gamma_{10} + \gamma_{11} Ed_i + U_{1i} \]
- Linear Slope for person \( i \)
- Fixed Linear Time Slope when Time=0 and Ed=12
- \( \Delta \) in Linear Time Slope per unit \( \Delta \) in Ed (=Ed*time)
- Random (Deviation) Linear Time Slope after controlling for Ed

\[ \beta_{2i} = \gamma_{20} + \gamma_{21} Ed_i + U_{2i} \]
- Quad Slope for person \( i \)
- Fixed Quad Time Slope when Ed = 12
- \( \Delta \) in Quad Time Slope per unit \( \Delta \) in Ed (=Ed*time^2)
- Random (Deviation) Quad Time Slope after controlling for Ed
Education (12 years = 0) as a Time-Invariant Predictor: Example using a Random Quadratic Time Model

Level 1: $y_{ti} = \beta_0 + \beta_1 Time_{ti} + \beta_2 Time_{ti}^2 + e_{ti}$

Level 2 Equations (one per $\beta$):

- $\beta_{0i} = \gamma_{00} + \gamma_{01} Ed_i + U_{0i}$
- $\beta_{1i} = \gamma_{10} + \gamma_{11} Ed_i + U_{1i}$
- $\beta_{2i} = \gamma_{20} + \gamma_{21} Ed_i + U_{2i}$

- Composite equation:

  $y_{ti} = (\gamma_{00} + \gamma_{01} Ed_i + U_{0i}) +$
  \[ (\gamma_{10} + \gamma_{11} Ed_i + U_{1i}) Time_{ti} + \]
  \[ (\gamma_{20} + \gamma_{21} Ed_i + U_{2i}) Time_{ti}^2 + e_{ti} \]

$\gamma_{11}$ and $\gamma_{21}$ are known as “cross-level” interactions (level-1 predictor by level-2 predictor)
**Fixed Effects of Time-Invariant Predictors**

- **Question of interest:** *Why do people change differently?*
  - We’re trying to predict individual differences in intercepts and slopes (i.e., reduce level-2 random effects variances)
  - So level-2 random effects variances become ‘conditional’ on predictors → actually random effects variances *left over*

\[
\begin{align*}
\beta_{0i} &= \gamma_{00} + U_{0i} & \beta_{0i} &= \gamma_{00} + \gamma_{01} Ed_i + U_{0i} \\
\beta_{1i} &= \gamma_{10} + U_{1i} & \beta_{1i} &= \gamma_{10} + \gamma_{11} Ed_i + U_{1i} \\
\beta_{2i} &= \gamma_{20} + U_{2i} & \beta_{2i} &= \gamma_{20} + \gamma_{21} Ed_i + U_{2i}
\end{align*}
\]

- Can calculate pseudo-$R^2$ for each level-2 random effect variance between models with *fewer* versus *more* parameters as:

\[
Pseudo R^2 = \frac{\text{random variance}_{\text{fewer}} - \text{random variance}_{\text{more}}}{\text{random variance}_{\text{fewer}}}
\]
Fixed Effects of Time-Invariant Predictors

• What about predicting level-1 effects with no random variance?
  - If the random linear time slope is n.s., can I test interactions with time?

  **This should be ok to do…**
  \[ \beta_{0i} = \gamma_{00} + \gamma_{01} Ed_i + U_{0i} \]
  \[ \beta_{1i} = \gamma_{10} + \gamma_{11} Ed_i + U_{1i} \]
  \[ \beta_{2i} = \gamma_{20} + \gamma_{21} Ed_i + U_{2i} \]

  **Is this still ok to do?**
  \[ \beta_{0i} = \gamma_{00} + \gamma_{01} Ed_i + U_{0i} \]
  \[ \beta_{1i} = \gamma_{10} + \gamma_{11} Ed_i \]
  \[ \beta_{2i} = \gamma_{20} + \gamma_{21} Ed_i \]

  - YES, surprisingly enough….
  - **In theory**, if a level-1 effect does not vary randomly over individuals, then it has “no” variance to predict (so cross-level interactions with that level-1 effect are not necessary)
  - However, because power to detect random effects is often lower than power to detect fixed effects, fixed effects of predictors can still be significant even if there is “no” \((\approx 0)\) variance for them to predict
  - Just make sure you test for random effects BEFORE testing any cross-level interactions with that level-1 predictor!
Let's say we have a significant fixed linear effect of time. What happens after we test a sex*time interaction?

<table>
<thead>
<tr>
<th>Random time slope</th>
<th>Non-Significant Sex*Time effect?</th>
<th>Significant Sex*Time effect?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random time slope initially <strong>not</strong> significant</td>
<td>Linear effect of time is <strong>FIXED</strong></td>
<td>Linear effect of time is systematically varying</td>
</tr>
<tr>
<td>Random time initially sig, <strong>not</strong> sig. after sex*time</td>
<td>---</td>
<td>Linear effect of time is systematically varying</td>
</tr>
<tr>
<td>Random time initially sig, <strong>still</strong> sig. after sex*time</td>
<td>Linear effect of time is <strong>RANDOM</strong></td>
<td>Linear effect of time is <strong>RANDOM</strong></td>
</tr>
</tbody>
</table>

The effects of level-1 predictors (time-level) can be fixed, random, or systematically varying. The effects of level-2 predictors (person-level) can only be fixed or systematically varying (nothing to be random over...yet).
Variance Accounted For By Level-2 Time-Invariant Predictors

• **Fixed effects of level 2 predictors *by themselves***:
  - L2 (BP) main effects (e.g., sex) reduce L2 (BP) random intercept variance
  - L2 (BP) interactions (e.g., sex by ed) also reduce L2 (BP) random intercept variance

• **Fixed effects of cross-level interactions (level 1* level 2):**
  - If the interacting level-1 predictor is random, any cross-level interaction with it will reduce its corresponding level-2 BP random slope variance
    - e.g., if *time* is random, then sex*time, ed*time, and sex*ed*time can each reduce the random linear time slope variance
  - If the interacting level-1 predictor not random, any cross-level interaction with it will reduce the level-1 WP residual variance instead
    - e.g., if *time* is fixed, then sex*time<sup>2</sup>, ed*time<sup>2</sup>, and sex*ed*time<sup>2</sup> will reduce the L1 (WP) residual variance → Different quadratic slopes from sex and ed will allow better trajectories, reduce the variance around trajectories
Variance Accounted for... For Real

- **Pseudo-$R^2$** is named that way for a reason... piles of variance can shift around, such that it can actually be negative
  - Sometimes a sign of model mis-specification
  - Hard to explain to readers when it happens!

- **One last simple alternative: Total $R^2$**
  - Generate model-predicted y’s from fixed effects only (NOT including random effects) and correlate with observed y’s
  - Then square correlation $\rightarrow$ total $R^2$
  - Total $R^2$ = total reduction in overall variance of y across levels
  - Can be “unfair” in models with large unexplained sources of variance

- **MORAL OF THE STORY:** Specify EXACTLY which kind of pseudo-$R^2$ you used—give the formula and the reference!!
The Joy of Time-Varying Predictors

- TV predictors predict leftover **WP (residual) variation:**

  - Modeling time-varying predictors is complicated because they represent an **aggregated effect:**
    - Effect of the *between-person* variation in the predictor $x_{ti}$ on $Y$
    - Effect of the *within-person* variation in the predictor $x_{ti}$ on $Y$
    - Here we are assuming the predictor $x_{ti}$ only **fluctuates** over time...
      - We will need a different model if $x_{ti}$ changes systematically over time...

[Diagram showing WP Change Model and WP Variation Model]
The Joy of Time-Varying Predictors

- Time-varying (TV) predictors usually carry 2 kinds of effects because they are really 2 predictor variables, not 1

- Example: Stress measured daily
  - Some days are worse than others:
    - WP variation in stress (represented as deviation from own mean)
  - Some people just have more stress than others all the time:
    - BP variation in stress (represented as person mean predictor over time)

- Can quantify each source of variation with an ICC
  - ICC = (BP variance) / (BP variance + WP variance)
  - ICC > 0? TV predictor has BP variation (so it could have a BP effect)
  - ICC < 1? TV predictor has WP variation (so it could have a WP effect)
Between-Person vs. Within-Person Effects

- Between-person and within-person effects in SAME direction
  - Stress → Health?
    - BP: People with more chronic stress than other people may have worse general health than people with less chronic stress
    - WP: People may feel worse than usual when they are currently under more stress than usual (regardless of what “usual” is)

- Between-person and within-person effects in OPPOSITE directions
  - Exercise → Blood pressure?
    - BP: People who exercise more often generally have lower blood pressure than people who are more sedentary
    - WP: During exercise, blood pressure is higher than during rest

- Variables have different meanings at different levels!
- Variables have different scales at different levels
3 Kinds of Effects for TV Predictors

- **Is the Between-Person (BP) effect significant?**
  - Are people with higher predictor values than other people (on average over time) also higher on Y than other people (on average over time), such that the person mean of the TV predictor accounts for level-2 random intercept variance ($\tau_{U_0}^2$)?

- **Is the Within-Person (WP) effect significant?**
  - If you have higher predictor values than usual (at this occasion), do you also have higher outcomes values than usual (at this occasion), such that the within-person deviation of the TV predictor accounts for level-1 residual variance ($\sigma_e^2$)?

- **Are the BP and WP effects different sizes: Is there a contextual effect?**
  - After controlling for the absolute value of TV predictor at each occasion, is there still an incremental contribution from having a higher person mean of the TV predictor (i.e., does one’s general tendency predict $\tau_{U_0}^2$ above and beyond)?
  - If there is no contextual effect, then the BP and WP effects of the TV predictor show convergence, such that their effects are of equivalent magnitude.
Modeling TV Predictors (labeled as $x_{ti}$)

- **Level-2 effect of $x_{ti}$:**
  - The level-2 effect of $x_{ti}$ is usually represented by the person’s mean of time-varying $x_{ti}$ across time (labeled as $PMx_i$ or $\bar{X}_i$)
  - $PMx_i$ should be centered at a **CONSTANT** (grand mean or other) so that 0 is meaningful, just like any other time-invariant predictor

- **Level-1 effect of $x_{ti}$ can be included two different ways:**
  - “**Group-mean-centering**” $\rightarrow$ “**person-mean-centering**” in longitudinal, in which level-1 predictors are centered using a level-2 **VARIABLE**
  - “**Grand-mean-centering**” $\rightarrow$ level-1 predictors are centered using a **CONSTANT** (not necessarily the grand mean; it’s just called that)
  - Note that these 2 choices do NOT apply to the level-2 effect of $x_{ti}$!
    - But the interpretation of the level-2 effect of $x_{ti}$ WILL DIFFER based on which centering method you choose for the level-1 effect of $x_{ti}$!
Person-Mean-Centering (P-MC)

• In P-MC, we decompose the TV predictor $x_{ti}$ into 2 variables that directly represent its BP (level-2) and WP (level-1) sources of variation, and include those variables as the predictors instead:

• **Level-2, PM predictor = person mean of $x_{ti}$**
  - $PMx_i = \bar{X}_i - C$
  - $PMx_i$ is centered at a constant $C$, chosen so 0 is meaningful
  - $PMx_i$ is positive? Above sample mean $\rightarrow$ “more than other people”
  - $PMx_i$ is negative? Below sample mean $\rightarrow$ “less than other people”

• **Level-1, WP predictor = deviation from person mean of $x_{ti}$**
  - $WPx_{ti} = x_{ti} - \bar{X}_i$ (note: uncentered person mean $\bar{X}_i$ is used to center $x_{ti}$)
  - $WPx_{ti}$ is NOT centered at a constant; is centered at a VARIABLE
  - $WPx_{ti}$ is positive? Above your own mean $\rightarrow$ “more than usual”
  - $WPx_{ti}$ is negative? Below your own mean $\rightarrow$ “less than usual”
Within-Person Fluctuation Model with Person-Mean-Centered Level-1 $x_{ti}$

$\rightarrow$ WP and BP Effects directly through separate parameters

$x_{ti}$ is person-mean-centered into $WPx_{ti}$, with $PMx_i$ at L2:

**Level 1:** $y_{ti} = \beta_0i + \beta_{1i}(WPx_{ti}) + e_{ti}$

**Level 2:** $\beta_0i = Y_{00} + Y_{01}(PMx_i) + U_{0i}$

$\beta_{1i} = Y_{10}$

- $Y_{10} = WP$ main effect of having more $x_{ti}$ than usual
- $Y_{01} = BP$ main effect of having more $\bar{X}_i$ than other people

$WPx_{ti} = x_{ti} - \bar{X}_i \rightarrow$ it has only Level-1 WP variation

$PMx_i = \bar{X}_i - C \rightarrow$ it has only Level-2 BP variation

Because $WPx_{ti}$ and $PMx_i$ are uncorrelated, each gets the total effect for its level (WP=L1, BP=L2)
Between-Person Effect = slope through person means = 1
Within-Person Effect = slope of individual lines = 0
Contextual Effect = difference of WP vs. BP slopes = +1

Person-Mean-Centered
Fixed Effects:
PMstress γ₀₁ = 1
WPstress γ₁₀ = 0
NO Between-Person Effect, ALL Within-Person Effect

- **Between-Person Effect** = slope through person means = 0
- **Within-Person Effect** = slope of individual lines = 1
- **Contextual Effect** = difference of WP vs. BP slopes = -1

Person-Mean-Centered

Fixed Effects:
- \( PMstress \gamma_{01} = 0 \)
- \( WPstress \gamma_{10} = 1 \)
Between-Person Effect $>$ Within-Person Effect

**Between-Person Effect** = slope through person means = 2

**Within-Person Effect** = slope of individual lines = 1

**Contextual Effect** = difference of WP vs. BP slopes = +1

**Person-Mean-Centered**

Fixed Effects:
- $PM_{stress} \gamma_{01} = 2$
- $WP_{stress} \gamma_{10} = 1$
Within-Person Fluctuation Model with Person-Mean-Centered Level-1 $x_{ti}$

$\rightarrow$ WP and BP Effects directly through separate parameters

$x_{ti}$ is person-mean-centered into $WPx_{ti}$, with $PMx_i$ at L2:

**Level 1:**

$$y_{ti} = \beta_{0i} + \beta_{1i}(WPx_{ti}) + e_{ti}$$

- $\beta_{0i} = Y_{00} + Y_{01}(PMx_i) + U_{0i}$
- $\beta_{1i} = Y_{10} + Y_{11}(PMx_i) + U_{1i}$

**Level 2:**

$$WPx_{ti} = x_{ti} - \overline{x}_i \rightarrow \text{it has only Level-1 WP variation}$$

$$PMx_i = \overline{x}_i - C \rightarrow \text{it has only Level-2 BP variation}$$

$U_{1i}$ is a random slope for the WP effect of $x_{ti}$

$Y_{10}$ = WP simple main effect of having more $x_{ti}$ than usual for $PMx_i = 0$

$Y_{01}$ = BP simple main effect of having more $\overline{x}_i$ than other people for people at their own mean ($WPx_{ti} = x_{ti} - \overline{x}_i \rightarrow 0$)

$Y_{11}$ = BP*WP interaction: how the effect of having more $x_{ti}$ than usual differs by how much $\overline{x}_i$ you have

Note: this model should also test $Y_{02}$ for $PMx_i * PMx_i$ (stay tuned)
Between-Person x Within-Person Interaction

**Between-Person Effect** = slope through person means = 2
**Within-Person Effect** = slope of individual lines = 1
**Contextual Effect** = difference of WP vs. BP slopes = +1

This model also includes a BP*WP interaction of −0.5, such that the within-person effect becomes weaker by 0.5 for every unit higher in mean stress.

**Person-Mean-Centered Fixed Effects:**
- **PMstress** $\gamma_{01} = 2$
- **WPstress** $\gamma_{10} = 1$
- **PM*WP** $\gamma_{10} = -0.5$

**Between-Person Effect** = slope through person means = 2
**Within-Person Effect** = slope of individual lines = 1
**Contextual Effect** = difference of WP vs. BP slopes = +1

Mean Stress = 4
Mean Stress = 5
Mean Stress = 6

Severity Outcome

Time-Varying Stress
3 Kinds of Effects for TV Predictors

• **What Person-Mean-Centering tells us directly:**

• **Is the Between-Person (BP) effect significant?**
  - Are people with higher predictor values than other people *(on average over time)* also higher on Y than other people *(on average over time)*, such that the person mean of the TV predictor accounts for level-2 random intercept variance *(τ_Y)^2*?
  - This would be indicated by a significant fixed effect of PMX_i
  - Note: this is NOT controlling for the absolute value of x_{ti} at each occasion

• **Is the Within-Person (WP) effect significant?**
  - If you have higher predictor values than usual *(at this occasion)*, do you also have higher outcomes values than usual *(at this occasion)*, such that the within-person deviation of the TV predictor accounts for level-1 residual variance *(σ_e)^2*?
  - This would be indicated by a significant fixed effect of WPx_{ti}
  - Note: this is represented by the relative value of x_{ti}, NOT the absolute value of x_{ti}
3 Kinds of Effects for TV Predictors

• What Person-Mean-Centering DOES NOT tell us directly:

• Are the BP and WP effects different sizes: Is there a contextual effect?
  ➢ After controlling for the absolute value of the TV predictor at each occasion, is there still an incremental contribution from having a higher person mean of the TV predictor (i.e., does one’s general tendency predict $\tau_{U0}^2$ above and beyond just the time-specific value of the predictor)?
  ➢ If there is no contextual effect, then the BP and WP effects of the TV predictor show convergence, such that their effects are of equivalent magnitude

• To answer this question about the contextual effect for the incremental contribution of the person mean, we have two options:
  ➢ Ask for the contextual effect via an ESTIMATE statement in SAS (or TEST in SPSS, or NEW in Mplus, or LINCOM in STATA): $WP_{xti} - 1$ $PM_{xi} 1$
  ➢ Use “grand-mean-centering” for time-varying $x_{ti}$ instead: $TVx_{ti} = x_{ti} - C$
  ➢ centered at a CONSTANT, NOT A LEVEL-2 VARIABLE
    ▪ Which constant only matters for what the reference point is; it could be the grand mean or other
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:
  – \( \beta_1 \) is \textbf{ALL the relationship} between \( X_{1i} \) and \( Y_i \)
  – \( \beta_2 \) is \textbf{ALL the relationship} between \( X_{2i} \) and \( Y_i \)

• If \( X_{1i} \) and \( X_{2i} \) ARE correlated:
  – \( \beta_1 \) is \textbf{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
  – \( \beta_2 \) is \textbf{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…
Person-MC vs. Grand-MC for Time-Varying Predictors

<table>
<thead>
<tr>
<th>Level 2</th>
<th>Original</th>
<th>Person-MC Level 1</th>
<th>Grand-MC Level 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{X}_i$</td>
<td>$PMx_i = \bar{X}_i - 5$</td>
<td>$x_{ti}$</td>
<td>$WPx_{ti} = x_{ti} - \bar{X}_i$</td>
</tr>
<tr>
<td>3</td>
<td>-2</td>
<td>2</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>-2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>6</td>
<td>-1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>

Same $PMx_i$ goes into the model using either way of centering the level-1 variable $x_{ti}$

Using Person-MC, $WPx_{ti}$ has NO level-2 BP variation, so it is not correlated with $PMx_i$

Using Grand-MC, $TVx_{ti}$ STILL has level-2 BP variation, so it is STILL CORRELATED with $PMx_i$

So the effects of $PMx_i$ and $TVx_{ti}$ when included together under Grand-MC will be different than their effects would be if they were by themselves...
Within-Person Fluctuation Model with $x_{ti}$ represented at Level 1 Only:

→ WP and BP Effects are **Smushed Together**

$x_{ti}$ is grand-mean-centered into $TVx_{ti}$, **WITHOUT** $PMx_i$ at L2:

**Level 1:** $y_{ti} = \beta_{0i} + \beta_{1i}(TVx_{ti}) + e_{ti}$

**Level 2:**

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

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

$\gamma_{10} = \text{*smushed* WP and BP effects}$

**TVx_{ti} = x_{ti} - c** → it still has both Level-2 BP and Level-1 WP variation

Because $TVx_{ti}$ still contains its original 2 different kinds of variation (BP and WP), 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 Time-Varying Predictor

Convergence Effect: \( \gamma_{\text{conv}} \approx \frac{\gamma_{BP}}{SE_{BP}^2} + \frac{\gamma_{WP}}{SE_{WP}^2} \)

- The convergence effect will often be closer to the within-person 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 time-varying predictor, convergence is testable by including a contextual effect (carried by the person mean) for how the BP effect differs from the WP effect...

Adapted from Raudenbush & Bryk (2002, p. 138)
Within-Person Fluctuation Model with Grand-Mean-Centered Level-1 $x_{ti}$

→ Model tests difference of WP vs. BP effects (It’s been fixed!)

$x_{ti}$ is grand-mean-centered into $TVx_{ti}$, WITH $PMx_i$ at L2:

**Level 1:** $y_{ti} = \beta_{0i} + \beta_{1i}(TVx_{ti}) + e_{ti}$

**Level 2:** $\beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i}$

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

- $\gamma_{10}$ becomes the WP effect → *unique* level-1 effect after controlling for $PMx_i$
- $\gamma_{01}$ becomes the contextual effect that indicates how the BP effect differs from the WP effect → *unique* level-2 effect after controlling for $TVx_{ti}$

$TVx_{ti} = x_{ti} - C \rightarrow$ it still has both Level-2 BP and Level-1 WP variation

$PMx_i = \bar{X}_i - C \rightarrow$ it has only Level-2 BP variation

→ does usual level matter beyond current level?
Person-MC and Grand-MC Models are Equivalent

Given a Fixed Level-1 Main Effect Only

**Person-MC:**  \[ WPx_{ti} = x_{ti} - PMx_i \]

Level-1:  \[ y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti} - PMx_i) + e_{ti} \]

Level-2:  \[ \beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i} \]

\[ \beta_{1i} = \gamma_{10} \]

\[ y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti} - PMx_i) + U_{0i} + e_{ti} \]

\[ y_{ti} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti} \]

**Grand-MC:**  \[ TVx_{ti} = x_{ti} \]

Level-1:  \[ y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti}) + e_{ti} \]

Level-2:  \[ \beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i} \]

\[ \beta_{1i} = \gamma_{10} \]

\[ y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti} \]

**Composite Model:**

- In terms of P-MC
- In terms of G-MC

<table>
<thead>
<tr>
<th>Effect</th>
<th>P-MC</th>
<th>G-MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>( \gamma_{00} )</td>
<td>( \gamma_{00} )</td>
</tr>
<tr>
<td>WP Effect</td>
<td>( \gamma_{10} )</td>
<td>( \gamma_{10} )</td>
</tr>
<tr>
<td>Contextual</td>
<td>( \gamma_{01} - \gamma_{10} )</td>
<td>( \gamma_{01} )</td>
</tr>
<tr>
<td>BP Effect</td>
<td>( \gamma_{01} )</td>
<td>( \gamma_{01} + \gamma_{10} )</td>
</tr>
</tbody>
</table>
P-MC vs. G-MC: Interpretation Example

Between-Person Effect = slope through person means = 2
Within-Person Effect = slope of individual lines = 0.5
Contextual Effect = difference of WP vs. BP slopes = +1.5

The contextual effect is given by the vertical distance along black line holding constant stress = 5.

Person-MC Fixed Effects:
- PMstress $\gamma_{01} = 2.0 = \text{BP}$
- WPstress $\gamma_{10} = 0.5 = \text{WP}$

Grand-MC Fixed Effects:
- PMstress $\gamma_{00} = 1.5 = \text{contextual}$
- TVstress $\gamma_{10} = 0.5 = \text{WP}$
Summary: 3 Effects for TV Predictors

- **Is the Between-Person (BP) effect significant?**
  - Are people with higher predictor values than other people (on average over time) also higher on Y than other people (on average over time), such that the person mean of the TV predictor accounts for level-2 random intercept variance ($\tau_{U0}^2$)?
  - Given directly by level-2 effect of $PMx_i$ if using Person-MC for the level-1 predictor (or can be requested via ESTIMATE if using Grand-MC for the level-1 predictor)

- **Is the Within-Person (WP) effect significant?**
  - If you have higher predictor values than usual (at this occasion), do you also have higher outcomes values than usual (at this occasion), such that the within-person deviation of the TV predictor accounts for level-1 residual variance ($\sigma_e^2$)?
  - Given directly by the level-1 effect of $WPx_{ti}$ if using Person-MC —OR — given directly by the level-1 effect of $TVx_{ti}$ if using Grand-MC and including $PMx_i$ at level 2 (without $PMx_i$, the level-1 effect of $TVx_{ti}$ if using Grand-MC is the smushed effect)

- **Are the BP and WP Effects different sizes: Is there a contextual effect?**
  - After controlling for the absolute value of TV predictor value at each occasion, is there still an incremental contribution from having a higher person mean of the TV predictor (i.e., does one’s general tendency predict $\tau_{U0}^2$ above and beyond)?
  - Given directly by level-2 effect of $PMx_i$ if using Grand-MC for the level-1 predictor (or can be requested via ESTIMATE if using Person-MC for the level-1 predictor)
Variance Accounted For By Level-1 Predictors

- **Fixed effects of level 1 predictors by themselves:**
  - Level-1 (WP) main effects reduce Level-1 (WP) residual variance
  - Level-1 (WP) interactions also reduce Level-1 (WP) 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., Person-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:
      \[
      \text{True } \tau_{U_0}^2 = \text{observed } \tau_{U_0}^2 - \frac{\sigma_e^2}{n} \rightarrow \text{so if only } \sigma_e^2 \text{ decreases, } \tau_{U_0}^2 \text{ increases}
      \]
The Joy of Interactions Involving Time-Varying Predictors

- Must consider interactions with both its BP and WP parts:
  - Example: Does time-varying stress ($x_{ti}$) interact with sex ($Sex_i$)?

- Person-Mean-Centering:
  - $WPx_{ti} \ast Sex_i \rightarrow$ Does the WP stress effect differ between men and women?
  - $PMx_i \ast Sex_i \rightarrow$ Does the BP stress effect differ between men and women?
    - Not controlling for current levels of stress
    - If forgotten, then $Sex_i$ moderates the stress effect only at level 1 (WP, not BP)

- Grand-Mean-Centering:
  - $TVx_{ti} \ast Sex_i \rightarrow$ Does the WP stress effect differ between men and women?
  - $PMx_i \ast Sex_i \rightarrow$ Does the *contextual* stress effect differ b/t men and women?
    - Incremental BP stress effect *after controlling for current levels of stress*
    - If forgotten, then although the level-1 main effect of stress has been un-smushed via the main effect of $PMx_i$, the interaction of $TVx_{ti} \ast Sex_i$ would still be smushed
Interactions with Time-Varying Predictors: Example: TV Stress ($x_{ti}$) by Gender ($Sex_i$)

**Person-MC:** \[ WPx_{ti} = x_{ti} - PMx_i \]

**Level-1:** \[ y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti} - PMx_i) + e_{ti} \]

**Level-2:** \[ \begin{align*}
\beta_{0i} &= \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PMx_i) + U_{0i} \\
\beta_{1i} &= \gamma_{10} + \gamma_{11}(Sex_i)
\end{align*} \]

**Composite:** \[ y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti} - PMx_i) + U_{0i} + e_{ti} + \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PMx_i) + \gamma_{11}(Sex_i)(x_{ti} - PMx_i) \]

**Grand-MC:** \[ TVx_{ti} = x_{ti} \]

**Level-1:** \[ y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti}) + e_{ti} \]

**Level-2:** \[ \begin{align*}
\beta_{0i} &= \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PMx_i) + U_{0i} \\
\beta_{1i} &= \gamma_{10} + \gamma_{11}(Sex_i)
\end{align*} \]

**Composite:** \[ y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti} + \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PMx_i) + \gamma_{11}(Sex_i)(x_{ti}) \]
Interactions Involving Time-Varying Predictors Belong at Both Levels of the Model

On the left below → Person-MC: \( WPx_{ti} = x_{ti} - PMx_i \)

\[
y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti} - PMx_i) + U_{0i} + e_{ti} \\
+ \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PMx_i) + \gamma_{11}(Sex_i)(x_{ti} - PMx_i)
\]

\[
y_{ti} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti} \\
+ \gamma_{02}(Sex_i) + (\gamma_{03} - \gamma_{11})(Sex_i)(PMx_i) + \gamma_{11}(Sex_i)(x_{ti})
\]

On the right below → Grand-MC: \( TVx_{ti} = x_{ti} \)

\[
y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti} \\
+ \gamma_{02}(Sex_i) + \gamma_{03}(Sex_i)(PMx_i) + \gamma_{11}(Sex_i)(x_{ti})
\]

 Intercept: \( \gamma_{00} = \gamma_{00} \)  
BP Effect: \( \gamma_{01} = \gamma_{01} + \gamma_{10} \)  
Contextual: \( \gamma_{01} = \gamma_{01} - \gamma_{10} \)

WP Effect: \( \gamma_{10} = \gamma_{10} \)  
BP*Sex Effect: \( \gamma_{03} = \gamma_{03} + \gamma_{11} \)  
Contextual*Sex: \( \gamma_{03} = \gamma_{03} - \gamma_{11} \)

Sex Effect: \( \gamma_{20} = \gamma_{20} \)  
BP*WP or Contextual*WP is the same: \( \gamma_{11} = \gamma_{11} \)

After adding an interaction for \( Sex_i \) with stress at both levels, then the Person-MC and Grand-MC models are equivalent.
Intra-variable Interactions

- Still must consider interactions with both its BP and WP parts!
- Example: Interaction of TV stress ($x_{ti}$) with person mean stress ($PM_{xi}$)

- **Person-Mean-Centering:**
  - $WP_{x_{ti}} * PM_{xi}$ → Does the WP stress effect differ by overall stress level?
  - $PM_{xi} * PM_{xi}$ → Does the BP stress effect differ by overall stress level?
    - Not controlling for current levels of stress
    - If forgotten, then $PM_{xi}$ moderates the stress effect only at level 1 (WP, not BP)

- **Grand-Mean-Centering:**
  - $TV_{x_{ti}} * PM_{xi}$ → Does the WP stress effect differ by overall stress level?
  - $PM_{xi} * PM_{xi}$ → Does the *contextual* stress effect differ by overall stress?
    - Incremental BP stress effect *after controlling for current levels of stress*
    - If forgotten, then although the level-1 main effect of stress has been un-smushed via the main effect of $PM_{xi}$, the interaction of $TV_{x_{ti}} * PM_{xi}$ would still be smushed
Intra-variable Interactions:
Example: TV Stress ($x_{ti}$) by Person Mean Stress ($PMx_i$)

**Person-MC:**
\[ WPx_{ti} = x_{ti} - PMx_i \]

**Level-1:**
\[ y_{ti} = \beta_0i + \beta_1i(x_{ti} - PMx_i) + e_{ti} \]

**Level-2:**
\[ \beta_0i = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{02}(PMx_i)(PMx_i) + U_{0i} \]
\[ \beta_1i = \gamma_{10} + \gamma_{11}(PMx_i) \]

**Composite:**
\[ y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti} - PMx_i) + U_{0i} + e_{ti} \]
\[ + \gamma_{02}(PMx_i)(PMx_i) + \gamma_{11}(PMx_i)(x_{ti} - PMx_i) \]

**Grand-MC:**
\[ TVx_{ti} = x_{ti} \]

**Level-1:**
\[ y_{ti} = \beta_0i + \beta_1i(x_{ti}) + e_{ti} \]

**Level-2:**
\[ \beta_0i = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{02}(PMx_i)(PMx_i) + U_{0i} \]
\[ \beta_1i = \gamma_{10} + \gamma_{11}(PMx_i) \]

**Composite:**
\[ y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti} \]
\[ + \gamma_{02}(PMx_i)(PMx_i) + \gamma_{11}(PMx_i)(x_{ti}) \]
Intra-variable Interactions:
Example: TV Stress \((x_{ti})\) by Person Mean Stress \((PM_{x_i})\)

On the left below → Person-MC: \(WP_{x_{ti}} = x_{ti} - PM_{x_i}\)

\[
y_{ti} = \gamma_{00} + \gamma_{01}(PM_{x_i}) + \gamma_{10}(x_{ti} - PM_{x_i}) + U_{0i} + e_{ti} \\
+ \gamma_{02}(PM_{x_i})(PM_{x_i}) + \gamma_{11}(PM_{x_i})(x_{ti} - PM_{x_i})
\]

On the right below → Grand-MC: \(TV_{x_{ti}} = x_{ti}\)

\[
y_{ti} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(PM_{x_i}) + \gamma_{10}(x_{ti}) + U_{0i} + e_{ti} \\
+ (\gamma_{02} - \gamma_{11})(PM_{x_i})(PM_{x_i}) + \gamma_{11}(PM_{x_i})(x_{ti})
\]

Intercept: \(\gamma_{00} = \gamma_{00}\)  
BP Effect: \(\gamma_{01} = \gamma_{01} + \gamma_{10}\)  
Contextual: \(\gamma_{01} = \gamma_{01} - \gamma_{10}\)

WP Effect: \(\gamma_{10} = \gamma_{10}\)  
BP² Effect: \(\gamma_{02} = \gamma_{02} + \gamma_{11}\)  
Contextual²: \(\gamma_{02} = \gamma_{02} - \gamma_{11}\)

BP*WP or Contextual*WP is the same: \(\gamma_{11} = \gamma_{11}\)

After adding an interaction for \(PM_{x_i}\) with stress at both levels, then the Person-MC and Grand-MC models are equivalent.
When Person-MC ≠ Grand-MC: Random Effects of TV Predictors

**Person-MC:**  \( WPx_{ti} = x_{ti} - PMx_i \)

**Level-1:**  \( y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti} - PMx_i) + e_{ti} \)

**Level-2:**  \( \beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i} \)

\( \beta_{1i} = \gamma_{10} + U_{1i} \)

\( \rightarrow y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti} - PMx_i) + U_{0i} + U_{1i}(x_{ti} - PMx_i) + e_{ti} \)

**Grand-MC:**  \( TVx_{ti} = x_{ti} \)

**Level-1:**  \( y_{ti} = \beta_{0i} + \beta_{1i}(x_{ti}) + e_{ti} \)

**Level-2:**  \( \beta_{0i} = \gamma_{00} + \gamma_{01}(PMx_i) + U_{0i} \)

\( \beta_{1i} = \gamma_{10} + U_{1i} \)

\( \rightarrow y_{ti} = \gamma_{00} + \gamma_{01}(PMx_i) + \gamma_{10}(x_{ti}) + U_{0i} + U_{1i}(x_{ti}) + e_{ti} \)

Variance due to \( PMx_i \) is removed from the random slope in Person-MC.

Variance due to \( PMx_i \) is still part of the random slope in Grand-MC. So these models cannot be made equivalent.
Random Effects of TV Predictors

- **Random intercepts** mean different things under each model:
  - **Person-MC** → Individual differences at \( WP_{x_{ti}} = 0 \) (that everyone has)
  - **Grand-MC** → Individual differences at \( TV_{x_{ti}} = 0 \) (that not everyone has)

- **Differential shrinkage of the random intercepts** results from differential reliability of the intercept data across models:
  - Person-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 too small** when using Grand-MC rather than Person-MC
  - Problem worsens with greater ICC of TV Predictor (more extrapolation)
  - Anecdotal example using clustered data was presented in Raudenbush & Bryk (2002; chapter 6)
Modeling Time-Varying Categorical Predictors

- Person-MC and Grand-MC really only apply to *continuous* TV predictors, but the need to consider BP and WP effects applies to *categorical* TV predictors too.

- Binary level-1 predictors do not lend themselves to Person-MC
  - e.g., \( x_{ti} = 0 \text{ or } 1 \) per occasion, person mean = .50 across occasions → impossible values
  - If \( x_{ti} = 0 \), then \( WPx_{ti} = 0 - .50 = -0.50 \); If \( x_{ti} = 1 \), then \( WPx_{ti} = 1 - .50 = 0.50 \)
  - Better: Leave \( x_{ti} \) uncentered and include person mean as level-2 predictor (results ~ Grand-MC)

- For >2 categories, person means of multiple dummy codes starts to break down, but we can think about types of people, and code BP effects accordingly

**Example: Dementia present/not at each time point?**

- **BP effects** → Ever diagnosed with dementia (no, yes)?
  - People who will eventually be diagnosed may differ prior to diagnosis (a BP effect)
- **TV effect** → Diagnosed with dementia at each time point (no, yes)?
  - Acute differences of before/after diagnosis logically can only exist in the “ever” people

- Other examples: Mentor status, father absence, type of shift work (AM/PM)
Wrapping Up: Person-MC vs. Grand-MC

• Time-varying predictors carry at least two potential effects:
  - Some people are higher/lower than other people $\rightarrow$ BP, level-2 effect
  - Some occasions are higher/lower than usual $\rightarrow$ WP, level-1 effect

• BP and WP effects almost always need to be represented by two or more model parameters, using either:
  - Person-mean-centering ($\text{WP}x_{ti}$ and $\text{PM}x_{ij}$): $\text{WP} \neq 0?$, $\text{BP} \neq 0?$
  - Grand-mean-centering ($\text{TV}x_{ti}$ and $\text{PM}x_{ij}$): $\text{WP} \neq 0?$, $\text{BP} \neq \text{WP}?$
  - Both yield equivalent models if the level-1 WP effect is fixed, but not if the level-1 WP effect is random
    - Grand MC $\rightarrow$ absolute effect of $x_{ti}$ varies randomly over people
    - Person MC $\rightarrow$ relative effect of $x_{ti}$ varies randomly over people
    - Use prior theory and empirical data (ML AIC, BIC) to decide
Checking for Violations of Model Assumptions: Why should we care?

• “Fitting a model with untenable assumptions is as senseless as fitting a model to data that are knowingly flawed” (Singer & Willett, pg. 127)

• HOWEVER:
  - We don’t actually know the true population relationships, so we don’t know when our estimates, SE’s, and p-values are off
  - Recommended strategy: “check assumptions of several initial models and any model you cite or interpret explicitly”
  - Mostly informal inspection – requires judgment call
    - “We prefer visual inspection of residual distributions” (S & W pg. 128)
  - Some things are fixable, some things are not
  - **End goal: Analyze the data the least wrong way possible** (because all models are wrong; some are useful)
General Consequences of Violating Model Assumptions

2 parts of the model to be concerned with:

- **Model for means = fixed effects**
  - Estimates depend on having the “right” model for the means → all relevant predictors, measured with as little error as possible
  - To the extent that predictors are missing or their effects are specified incorrectly, **fixed effect estimates will be biased**

- **Model for the variances = random effects and residuals**
  - SE and \( p \)-values of fixed effects depend on having the “right” model for the variances → most closely approximate actual data
  - To the extent that the model for the variances is off, **fixed effects SE and \( p \)-values will be off, too (biased)**
  - Because the general linear mixed model is estimated using a multivariate normal distribution for the \( V \) matrix, certain assumptions follow...
General Linear Mixed Model
Assumptions

- **GLM Assumptions:**
  - Normality of *residuals* (not outcomes)
  - Independence and constant variance of *residuals*
    - Across sampling units
    - Across predictors

- **MLM Assumptions are the same, except:**
  - Apply *at* each level and *across* levels
  - More general options are available for changing the model to accommodate violations of assumptions if needed (goal is to *transform the model*, not the data)
  - ML also assumes MAR for any missing outcomes
Plots to Assess Assumptions:

**Normality**
- **Histogram:** Normality of Residuals Does Not Hold
  - Positively Skewed Variable
  - Histogram: Positively Skewed Variable
  - Normal Probability Plot: Normality of Residuals Does Not Hold
  - **Positive Skew**

**Independence & Constant Variance**
- **Full Model:** +A, +B, +A*B
  - Flat and Even
  - Random Intercepts
  - Incomplete Model: +A, +B, -A*B
  - Flat, but Not Even
  - Random Intercepts
- **Incomplete Model:** -A, +B, +A*B
  - Not Flat, but Even
  - Random Intercepts
  - Incomplete Model: -A, +B, -A*B
  - Not Flat, Not Even
  - Random Intercepts
MLM Assumptions: Normality
Multiple ‘residuals’ to consider:

**Level-1** $e_{ti}$ residuals $\rightarrow$ (multivariate) normal distribution

- $e_{ti} \sim N(0, R)$ where $R = \sigma_e^2$
- $e_{ti}$ has a mean = 0 and some estimated variance(s) and potentially covariances as well (is an empirical question)

**Level-2** $U_i$’s $\rightarrow$ multivariate normal distribution

- $U_0i, U_{1i}, ... \sim N(0, G)$
- If random intercept: $G = \begin{pmatrix} \tau_{U0}^2 \end{pmatrix}$
- If random slopes: $G = \begin{pmatrix} \tau_{U0}^2 & \tau_{U01} \\ \tau_{U01} & \tau_{U1}^2 \end{pmatrix}$

- $U$’s EACH have a mean = 0 and some estimated variance, with estimated covariances between them
  - The actual mean of $U$ has another name: __________
  - Covariances not included by default: added with TYPE=UN
1. Pick a new model for the level-1 $e_{ti}$ residuals
   - **Generalized linear mixed models** to the rescue!
     - Binary $\rightarrow$ Logit or Probit, Ordinal $\rightarrow$ Cumulative Logit
     - Count $\rightarrow$ Poisson or Negative Binomial (+ Zero-Inflated versions)
   - Unfortunately, level-2 U’s are still assumed multivariate normal
     - Problems with skewness $\rightarrow$ random effects CI’s go out of bounds
   - Tricky to estimate, but should use ML with numeric integration when possible (try to avoid older “pseudo” or “quasi” ML options)

2. Transform your data... carefully if at all...
   - Assumptions apply to residuals, not to data!
   - Complicates interpretations (linear relationships $\rightarrow$ nonlinear)
   - Inherently subjective (especially “outlier” removal)
     - Check for extreme leverage on solution instead via INFLUENCE options after / on MODEL statement in PROC MIXED
3. Robust ML for Non-Normality

- **MLR in Mplus**: \( \approx \) Yuan-Bentler \( T_2 \) (permits MCAR or MAR missing)
  - Same estimates and -2LL, corrected standard errors for all model parameters

- **\( \chi^2 \)-based fit statistics** are adjusted based on an estimated **scaling factor**:
  - Scaling factor = 1.000 = perfectly multivariate normal = same as ML
  - Scaling factor > 1.000 = leptokurtosis (too-fat tails; fixes too big \( \chi^2 \))
  - Scaling factor < 1.000 = platykurtosis (too-thin tails; fixes too small \( \chi^2 \))

- **SEs** computed with Huber-White ‘sandwich’ estimator \( \rightarrow \) uses an information matrix from the variance of the partial first derivatives to correct the information matrix from the partial second derivatives
  - Leptokurtosis (too-fat tails) \( \rightarrow \) increases information; fixes too small SEs
  - Platykurtosis (too-thin tails) \( \rightarrow \) lowers information; fixes too big SEs

- **In SAS**: use “EMPIRICAL” option in PROC MIXED line
  - SEs are computed the same way but for fixed effects only, but can be unstable in unbalanced data, especially in small samples
  - SAS does not provide the needed scaling factor to adjust -2\( \Delta \)LL test
    (not sure if this is a problem if you just use the fixed effect p-values)
Independence of Residuals
At Level 1:

- **Level-1 e_{ti} residuals are uncorrelated across level-1 units**
  - Once random effects are modeled, residuals of the occasions from the same person are no longer correlated

- **Solution for clustered or longitudinal models:**
  - Choose the ‘right’ specification of random effects
    - Random effects go in \( G \); **what’s left in R is uncorrelated across observations**

- **Another solution for longitudinal models:**
  - Choose the ‘right’ alternative for the structure of the residual variances and covariances over time
  - Use \( R \) matrix or \( G \) and \( R \) matrices to better approximate observed data:
    - Are the residuals still correlated (AR1, TOEP) after random effects?
    - Are the variances over time homogeneous or heterogeneous?
      - This falls under the “constant variance” assumption – more on that later
Independence of Residuals
At Level 2:

- **Level-2** $U_i$'s are uncorrelated across level-2 units
  - Implies no additional effects of clustering/nesting across persons after controlling for person-level predictors

- **Two alternatives to deal with additional clustering/nesting:**
  - Via fixed effects: Add dummy codes as level-2 predictors
    - Adjusts model for mean differences, but DOES NOT allow you to predict those mean differences
  - Via random effects: Add more levels (e.g., for family, group)
    - Adjusts model for mean differences, and it DOES allow you to predict those mean differences
    - Like adding another part to $G$
Independence of Residuals Across Levels:

- Level-1 $e_{ti}$ residuals and Level-2 $U_i$’s are also uncorrelated
  - Implies that what’s left over at level-1 is not related to what’s left over at level 2
  - Could be violated if level-2 effects are not modeled separately from level-1 effects (i.e., if convergence of level-1 predictors is assumed when it shouldn’t be)

- Solution: Don’t smush anything!
  - Allow different effects across upper levels for any lower-level predictor with respect to both main effects and interactions
Constant Variance of Residuals Across Sampling Units:

- **Level-2 $U_i$’s** have constant variance across **level-2 units**
  - Implies no subgroups of individuals or groups that are more or less variable in terms of their distributions of random effects
  - If not, we can fit a heterogeneous variance model instead (stay tuned)

- **Level-1 $e_{ti}$ residuals** have constant variance across **level-2 units**
  - Implies equal unexplained within-person variability across persons
  - Check for missing random effects of level-1 X’s or cross-level interactions
  - If not, we can fit a heterogeneous variance model instead (stay tuned)

- **Level-1 $e_{ti}$ residuals** have constant variance across **level-1 units**
  - Implies equal unexplained within-person variability across occasions
  - Can add additional random slopes for time or fit a heterogeneous variance model instead (e.g., TOEPH instead of TOEP, data permitting)

- * Test for heterogeneity of level-1 residuals applicable sometimes if $n > 10$ or so (see Snijders & Bosker, 1999, p. 126-7)
Independence and Constant Variance of Residuals Across Predictors:

- **Level-1 $e_{ti}$ residuals** are flat with constant variance across **level-1 X’s**
  - Implies no remaining relationship of X-Y at **level 1**
  - Specific example: level-1 residuals are flat and even across time after fixed and random effects (but we can fit separate variances by time if needed)
  - Check for potential nonlinear effects of level-1 predictors

- **Level-2 $U_i$’s** are flat with constant variance across **level-1 X’s**
  - Only possible relation between level-2 $U_i$ and level-1 X is through relationship between level-2 PMx and level-2 $U_i$ (so include PMx to avoid smushing)

- **Level-1 $e_{ti}$ residuals** are flat with constant variance across **level-2 X’s**
  - If not, we can fit a heterogeneous variance model instead (stay tuned)

- **Level-2 $U_i$’s** are flat with constant variance across **level-2 X’s**
  - Implies no remaining relationship of X-Y at **level 2**
  - Check for potential nonlinear effects of level-2 predictors
  - If not, we can fit a heterogeneous variance model instead (stay tuned)
Heterogeneous Variance Models

- Besides having random effects, predictors can play a role in predicting heterogeneity of variance at their level or lower:
  - Level-2 predictors → Differential level-2 random effects variances $\tau^2_U$
  - Differential level-1 residual variances $\sigma^2_e$
  - Level-1 predictors → Differential level-1 residual variances $\sigma^2_e$
  - -2ΔLL tests used to see if extra heterogeneity effects are helpful

- **Level-2 predictor of level-2 random effects variances for WP change:**
  - e.g., changes in height over time in boys and in girls?
  - Boys may be taller and grow faster than girls on average
    - Effect of sex and sex*time → predict level of Y in **model for the means**
  - Boys may be more variable than girls in their levels and rates of change in height
    - Effect of sex → different $\tau^2_U$ in **level-2 model for the variances**
Heterogeneous Variance Models

• Level-2 predictor of level-2 and level-1 variances for WP fluctuation:
  ➢ e.g., daily fluctuation in mood in men and in women
  ➢ Men may have worse negative mood than women on average
    ▪ Effect of sex → predict level of Y in model for the means
  ➢ There may be greater variability among men than women in mean mood
    ▪ Effect of sex → different \( \tau_{ij}^2 \) in level-2 model for the variances
  ➢ Men may be more variable than women in their daily mood fluctuation
    ▪ Effect of sex → different \( \sigma_e^2 \) in level-1 model for the variances

• Level-1 predictor of level-1 variance for WP fluctuation:
  ➢ e.g., daily fluctuation in mood on stress/non-stress days
  ➢ Negative mood may be worse on average on stress days than non-stress days
    ▪ Effect of stress → predict level of Y in model for the means
  ➢ There may be greater variation in mood on stress days than on non-stress days
    ▪ Effect of stress → different \( \sigma_e^2 \) in level-1 model for the variances
Estimating Heterogeneous Variance Models via PROC MIXED

• Different variances via `GROUP=groupvar` option after the `/` on the
  RANDOM statement for level 2 or REPEATED statement for level 1
  - Less flexible than multiple-group SEM because the whole \( G \) and/or \( R \) matrix is
    either the same or different across groups (all or nothing)
  - `GROUP=` is limited to categorical predictors (must use CLASS statement)
    ▪ Continuous level-2 predictors must use NLMIXED custom function instead

• In addition, different level-1 residual variances can be modeled via the
  `LOCAL=EXP( )` option after `/` on REPEATED statement
  - For categorical or continuous level-2 or level-1 predictors
  - Cannot be used with any other \( R \) matrix structure besides VC
  - Predicts natural log of the residual variance so prediction can’t go negative:
    \[
    \sigma^2_{e,i} = \alpha_0 \left( \exp \left( \alpha_1 X_1 + \alpha_2 X_2 \right) \right)
    \]
Estimating Heterogeneous Variance Models via PROC NLMIXED

- Can also write custom variance functions (see Hedeker’s examples)
  - More flexible, linear models approach can accommodate any combination of categorical or continuous predictors
  - Here, an example of heterogeneous level-2 random intercept variance from Hoffman chapter 7 (see example for NLMIXED code)

**Level 1:**
Symptoms\(_{ti} = \beta_{0i} + e_{ti}\)
Residual Variance: \(\sigma_{e_{ti}}^2 = \exp[\eta_{0i}]\)

**Level 2:**
Intercept: \(\beta_{0i} = \gamma_{00} + \gamma_{01}(Women_i) + \gamma_{02}(Age_i - 80) + \gamma_{03}(Women_i)(Age_i - 80) + U_{0i}\)
Random Intercept Variance \(\tau_{U_{0i}}^2 = \exp[\nu_{00} + \nu_{01}(Women_i) + \nu_{02}(Age_i - 80) + \nu_{03}(Women_i)(Age_i - 80)]\)
Residual Variance: \(\eta_{0i} = \varepsilon_{00}\)

\(\nu\) are effects for differential random intercept variance by intercept, sex, age and sex by age
\(\eta_{0i}\) is a placeholder (like \(\beta\)’s in model for means)
\(\varepsilon_{00}\) is like fixed intercept of residual variance
Estimating Heterogeneous Variance Models via PROC NLMIXED

- Can test for a \( \omega \) “scale factor”—like a random intercept for individual differences in residual variance (in WP variation)

**Level 1:**
- Symptoms\( _{ti} = \beta_{0i} + e_{ti} \)
- Residual Variance: \( \sigma_{e_{ti}}^2 = \exp[\eta_{0i}] \)

**Level 2:**
- Intercept: \( \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Women}_i) + \gamma_{02}(\text{Age}_i - 80) + \gamma_{03}(\text{Women}_i)(\text{Age}_i - 80) + U_{0i} \)

Random Intercept Variance: \( \tau_{U_{0i}}^2 = \exp[\nu_{00}] \)

Residual Variance: \( \eta_{0i} = \varepsilon_{00} + \omega_{0i} \)

From Hoffman chapter 7 (see example for NLMIXED code)

No \( \nu \) predictors of differential random intercept variance, just an intercept here

\( \eta_{0i} \) is a placeholder (like \( \beta \)’s in model for means)
\( \varepsilon_{00} \) is like fixed intercept of residual variance
\( \omega_{0i} \) is like random intercept of residual variance
Estimating Heterogeneous Variance Models via PROC NLMIXED

Level 1:
Symptoms_{ti} = \beta_{0i} + \beta_{li} \left( \text{Mood}_{ti} - \overline{\text{Mood}}_{i} \right) + \beta_{2i} \left( \text{Stressor}_{ti} \right) + \epsilon_{ti}

Residual Variance: \sigma^2_{\epsilon_i} = \exp \left[ \eta_{0i} + \eta_{li} \left( \text{Mood}_{ti} - \overline{\text{Mood}}_{i} \right) + \eta_{2i} \left( \text{Stressor}_{ti} \right) \right]

Level 2:
Intercept: \beta_{0i} = \gamma_{00} + \gamma_{01} \left( \text{Women}_{i} \right) + \gamma_{02} \left( \text{Age}_{i} - 80 \right) + \gamma_{03} \left( \text{Women}_{i} \right) \left( \text{Age}_{i} - 80 \right)
+ \gamma_{04} \left( \overline{\text{Mood}}_{i} - 2 \right) + \gamma_{08} \left( \overline{\text{Stressor}}_{i} - 0.40 \right) + \gamma_{09} \left( \text{Women}_{i} \right) \left( \overline{\text{Stressor}}_{i} - 0.40 \right)
+ \gamma_{0,16} \left( \overline{\text{Mood}}_{i} - 2 \right)^2 + \upsilon_{0i}

Within-Person Mood: \beta_{li} = \gamma_{10} + \gamma_{14} \left( \overline{\text{Mood}}_{i} - 2 \right)

Within-Person Stressor: \beta_{2i} = \gamma_{20} + \gamma_{21} \left( \text{Women}_{i} \right)

Random Intercept Variance \tau^2_{U_{0i}} = \exp \left[ \upsilon_{00} + \upsilon_{01} \left( \text{Women}_{i} \right) + \upsilon_{02} \left( \text{Age}_{i} - 80 \right) + \upsilon_{04} \left( \overline{\text{Mood}}_{i} - 2 \right) + \upsilon_{08} \left( \overline{\text{Stressor}}_{i} - 0.40 \right) \right]

Residual Variance:
\eta_{0i} = \epsilon_{00} + \epsilon_{01} \left( \text{Women}_{i} \right) + \epsilon_{02} \left( \text{Age}_{i} - 80 \right) + \epsilon_{04} \left( \overline{\text{Mood}}_{i} - 2 \right) + \epsilon_{08} \left( \overline{\text{Stressor}}_{i} - 0.40 \right)
\eta_{li} = \epsilon_{10}
\eta_{2i} = \epsilon_{20}

\epsilon \text{ are predictors of differential residual variance}
\omega_{0i} \text{ was not estimable, so was not included}

From Hoffman chapter 8 (see example for NLMIXED code)

\upsilon \text{ predictors of differential random intercept variance}
Assumptions of MLM: Summary

• Because model estimates, SEs, and fit statistics are derived from likelihood estimation using the multivariate normal distribution, their accuracy depends on its assumptions being met:
  
  - Residuals at each level (level 1 = \( e_{ti} \) values, level 2 = \( U_i \) values) are
    
    1. normally distributed,
    2. uncorrelated at each level and across levels, 
    3. equally distributed across \( X \)’s at each level and across levels.

• If not:
  
  1. transform the data (meh) or pick a generalized model for non-linear outcomes 
     (better when possible), or use robust ML for corrected SE’s
  2. add fixed or random effects (or a correlation over time),
  3. make sure predictive relationships are correctly specified, and then consider heterogeneous variance models if needed.