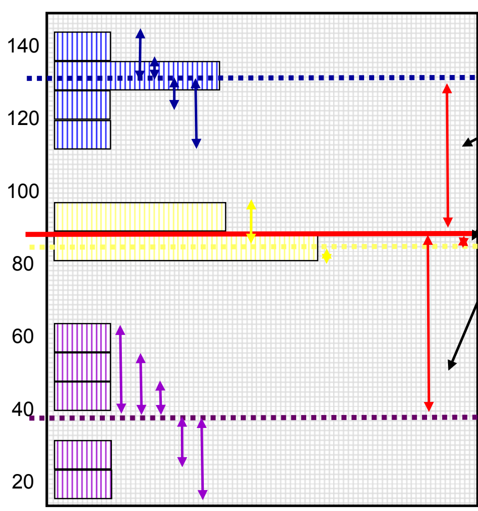
Lecture 2: Concepts in Modeling Within-Person Change and Fluctuation

- **Empty Models and Intraclass Correlation**
- Fixed vs. Random Effects of Persons and Time
- Considering Alternative Metrics of Time
- Within-Person Change vs. Within-Person Fluctuation

Empty Longitudinal Multilevel Model: Review of Terminology



Variance of Y → 2 sources:

<u>Level 2 Random Intercept Variance</u> (of U_{0i}):

→ Between-Person Variance (τ_{U0}^2)

→ Difference from GRAND mean

→ INTER-Individual Differences

<u>Level 1 Residual Variance</u> (of e_{ti}):

→ Within-Person Variance (σ_e^2)

→ Difference from OWN mean

→ INTRA-Individual Differences

Empty* Multilevel Model

Model for the Means; Model for the Variance

General Linear Model

$$y_i = \beta_0 + e_i$$

Multilevel Model

Level 1: $y_{ti} = \beta_{0i} + e_{ti}$

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

Sample **Grand Mean Intercept**

Individual **Intercept Deviation**

3 Model Parameters

1 Fixed Effect:

γ_{00} → fixed intercept

1 Random Effect (intercept):

U_{0i} → person-specific deviation

→ mean=0, variance = τ_{U0}^2

1 Residual Error:

e_{ti} → time-specific deviation

→ mean=0, variance = σ_e^2

Composite equation: $y_{ti} = \gamma_{00} + U_{0i} + e_{ti}$

* To be more clear, I call this an "empty means, random intercept" model

Empty Multilevel Model: Useful Descriptive Statistic → ICC

Intraclass Correlation (ICC):

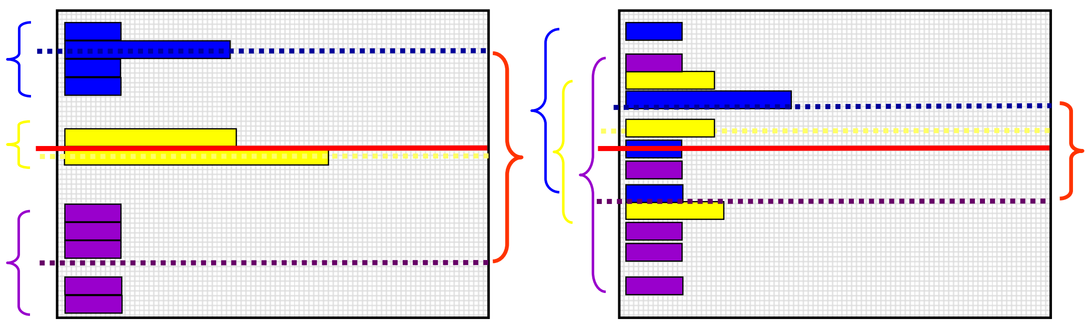
$$ICC = \frac{\text{Intercept Variance}}{\text{Intercept Variance + Residual Variance}} = \frac{\tau_{U_0}^2}{\tau_{U_0}^2 + \sigma_e^2}$$

$$ICC = \frac{\text{Between-Person Variance}}{\text{Between-Person Variance + Within-Person Variance}}$$

- ICC = Proportion of total variance that is between persons
- ICC = Average correlation among occasions
- ICC is a standardized way of expressing how much we need to worry about *dependency due to person mean differences* **(i.e., ICC is an effect size for *constant* person dependency)**

$$ICC = \frac{\text{Between-Person Variance}}{\text{Between-Person Variance + Within-Person Variance}}$$

Counter-Intuitive: Between-Person Variance is in the numerator, but the ICC is the correlation over time!



ICC = **BTW** / **BTW** + within

→ Large ICC

→ Large correlation over time

ICC = btw / btw + **WITHIN**

→ Small ICC

→ Small correlation over time

More New Vocabulary: Labels for 3 Types of Models

- "Empty Model" → "Empty means, random intercept model"
  - Just fixed intercept, random intercept variance, residual variance
  - First baseline model for everything to follow
  - Used to compute an ICC
- **"Unconditional (Growth) Model" is up next**
  - Effects related to time, but no other predictors yet
  - Second baseline model for everything to follow
- "Conditional (Growth) Model" is coming later
  - With predictors besides time
  - May end up with more than 1 depending on research questions

Lecture 2: Concepts in Modeling Within-Person Change and Fluctuation

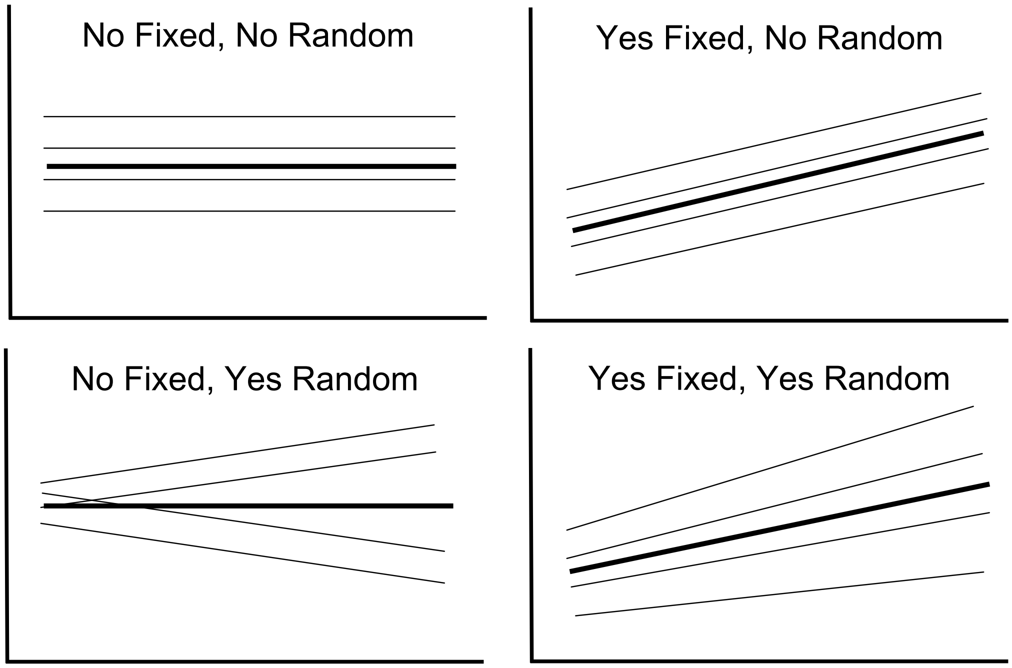
- Empty Models and Intraclass Correlation
- **Fixed vs. Random Effects of Persons and Time**
- Considering Alternative Metrics of Time
- Within-Person Change vs. Within-Person Fluctuation

Augmenting the Empty Model: 2 Questions for Effects of Time

1. Is there an effect of time **on average**?
   → Is the average line not flat?
   → Significant **Fixed** Effect of Time

2. Does the average effect of time **vary across individuals**?
   → Does each person need his/her own line?
   → Significant **Random** Effect of Time

Fixed and Random Effects of **Time**



Random Linear Time Model

6 Model Parameters

2 Fixed Effects: γ_{00} and γ_{10}

2 Random Effects (+1 covariance):
Variances of U_{0i} and U_{1i} (τ_{U0}^2, τ_{U1}^2)
Covariance of U_{0i} and U_{1i} (τ_{U01})

1 Residual Variance of e_{ti} (σ_e^2)

Multilevel Model

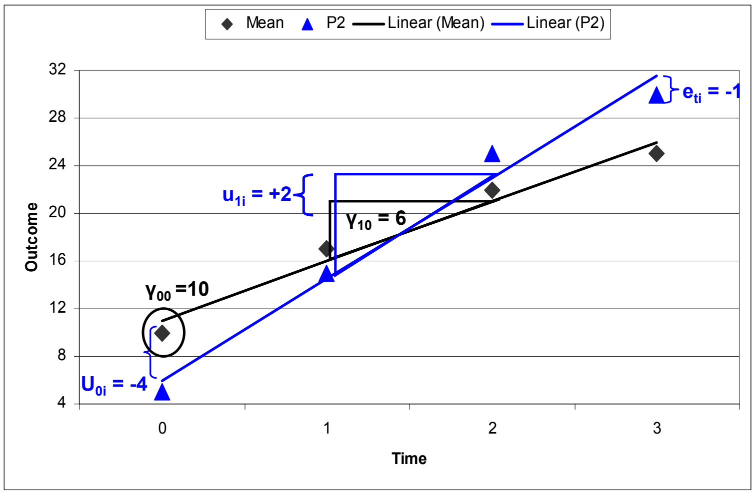
Level 1: $$y_{ti} = \beta_{0i} + \beta_{1i}Time_{ti} + e_{ti}$$

Level 2:
$$\beta_{0i} = \underbrace{\gamma_{00}}_{\text{Sample (Grand Mean) Intercept}} + \underbrace{U_{0i}}_{\text{Individual Intercept Deviation}} \qquad \beta_{1i} = \underbrace{\gamma_{10}}_{\text{Sample (Grand Mean) Slope}} + \underbrace{U_{1i}}_{\text{Individual Slope Deviation}}$$

Composite Model
$$y_{ti} = \underbrace{(\gamma_{00} + U_{0i})}_{\beta_{0i}} + \underbrace{(\gamma_{10} + U_{1i})}_{\beta_{1i}}Time_{ti} + e_{ti}$$

Fixed & Random Effects of Time

$$y_{ti} = (\underbrace{\gamma_{00}}_{\text{Fixed Intercept}} + \underbrace{U_{0i}}_{\text{Random Intercept Deviation}}) + (\underbrace{\gamma_{10}}_{\text{Fixed Slope}} + \underbrace{U_{1i}}_{\text{Random Slope Deviation}})Time_{ti} + \underbrace{e_{ti}}_{\text{error for person i at time t}}$$



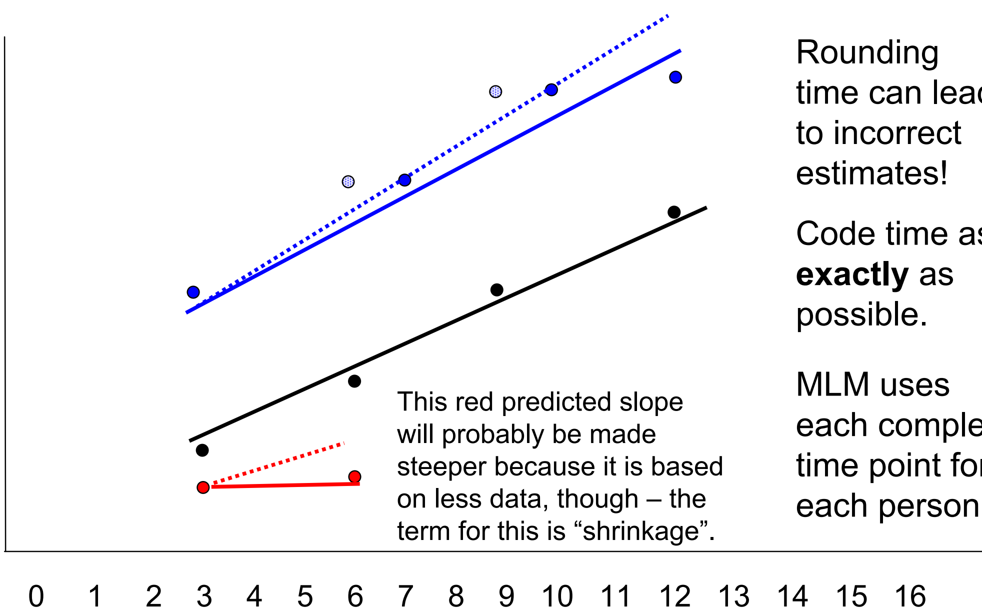
6 Model Parameters:

2 Fixed Effects:
γ_{00} Intercept, γ_{10} Slope

2(+1) Random Effects:
U_{0i} Intercept Variance
U_{1i} Slope Variance
Int-Slope Covariance

1 e_{ti} Residual Variance

Unbalanced Time → Different time measurements across persons? OK



Rounding time can lead to incorrect estimates!

Code time as **exactly** as possible.

MLM uses each complete time point for each person.

Longitudinal Data: Modeling Means and Variances

- We have two tasks in describing the effects of "time":

1. Choose a Model for the Means

- What kind of change in the outcome do we have on average?
- What kind of and how many parameters do we need to represent that change as parsimoniously but accurately as possible?

2. Choose a Model for the Variances

- What kind of pattern do the variances and covariances of the outcome show over time *due to individual differences in change*?
- What kind of and how many parameters do we need to represent that pattern as parsimoniously but accurately as possible?

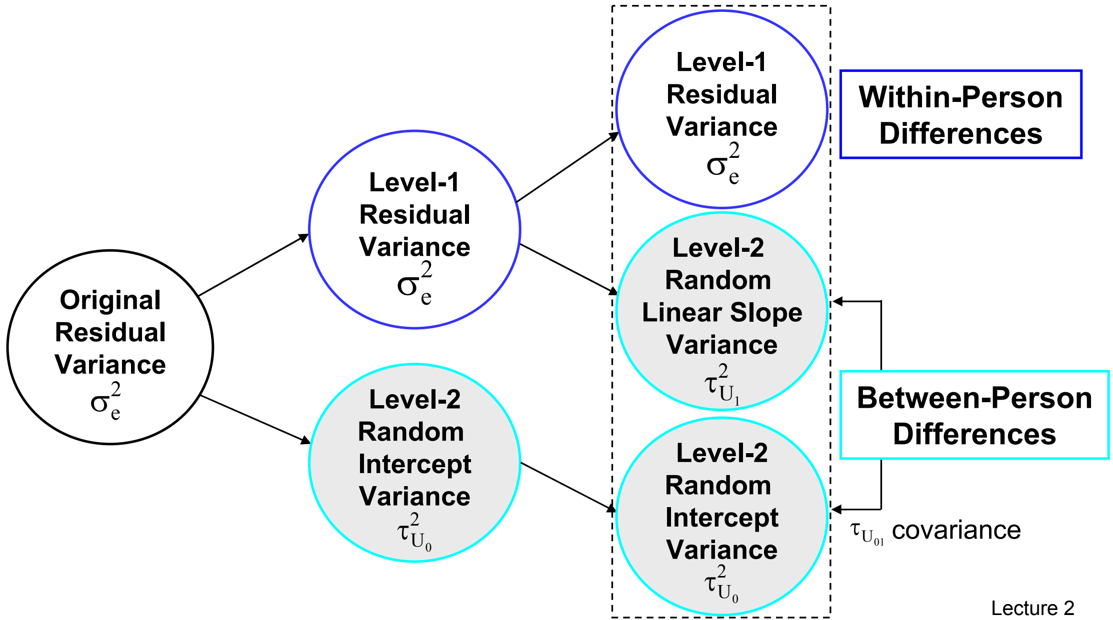
The Point of MLM: Dependency

- Common description of the purpose of MLM is that it 'addresses' or 'handles' correlated (dependent) data…
- But where does this 'correlation' come from?
  3 places (here, an example with health as an outcome):

1. *Mean differences across persons*
   - Some people are just healthier than others (at every time point)
2. *Differences in effects of predictors across persons*
   - Does *time* affect health more in some persons than others?
   - Does *daily stress* affect health more in some persons than others?
3. Non-constant within-person correlation for unknown reasons
   - Occasions closer together may just be more related
   - More likely for outcomes that fluctuate (than change) over time

2-Level Models for the Variances

- **Where does the correlation or 'dependency' go?**
  Into a new random effects variance component ('pile of variance')



The Point of MLM: Dependency

- Common description of the purpose of MLM is that it 'addresses' or 'handles' correlated (dependent) data…
  - In reality, the way that MLM 'handles' dependency is to keep making new piles of variance until each one is independent
    - "Piles of variance" = "variance components" = "random effects"
    - Piles of variance (i.e., random intercepts and slopes) are allowed to be correlated at the same level, but not across levels

- However, there are really two options for dealing with dependency:
  - As a *fixed* effect → goes into model for *means*
  - As a *random* effect → goes into model for *variances*
    - Random effects are really person*something interaction terms

Fixed vs. Random Effects in Longitudinal Models

- Why model effects of person (main effects, interactions with time) as random effects instead of fixed effects?
  - Could have used **fixed effects** in model for means instead:
    - Put in N-1 dummy codes for person → main effects of person
    - Put in N-1*time dummy codes per person → person*time interactions
    - This would control for dependency, and is what is actually used in some programs that "control" SEs for sampling/nesting
  - Important advantages of estimating **random effects** instead:
    - <u>Quantification</u>: Direct measure of how much of the outcome variance is due to person differences (in outcome level or effects of predictors)
    - <u>Prediction</u>: Person differences (main effects and effects of time) then become predictable quantities – this can't happen using fixed effects
  - **Summary: Random effects give you "predictable control"**

What the two different models for the variances actually imply…

- **<u>Random intercept only:</u>**
  - People differ from each other systematically in ONE way only
  - That ONE way is in intercept (U_{0i}), and that's it
  - If this is true, once you take out the individual differences in level (the U_{0i}'s), what's left (the e_{ti}'s) should be unrelated across persons as well as unrelated and homogeneous across time

- **<u>Random intercepts and slopes:</u>**
  - People differ from each other systematically in TWO ways
  - Those TWO ways are intercept (U_{0i}) and slope (U_{1i})
  - If this is true, if you only take out individual differences in level (the U_{0i}'s), what's left (the e_{ti}'s) will still have some correlation
  - Once you take out individual differences in slopes (the U_{1i}'s), what's left (the e_{ti}'s) should now be unrelated across persons as well as unrelated and homogeneous across time

Fixed Linear Time, Random Intercept Model

$$y_{ti} = \gamma_{00} + U_{0i} + (\gamma_{10}Time_{ti}) + e_{ti}$$

n = 4 occasions

Total Predicted Data Matrix is called V Matrix

$$\begin{bmatrix} \hat{\sigma}^2_{11} & \hat{\sigma}_{12} & \hat{\sigma}_{13} & \hat{\sigma}_{14} \\ \hat{\sigma}_{21} & \hat{\sigma}^2_{22} & \hat{\sigma}_{23} & \hat{\sigma}_{24} \\ \hat{\sigma}_{31} & \hat{\sigma}_{32} & \hat{\sigma}^2_{33} & \hat{\sigma}_{43} \\ \hat{\sigma}_{41} & \hat{\sigma}_{42} & \hat{\sigma}_{43} & \hat{\sigma}^2_{44} \end{bmatrix}$$

Level 2, BP Variance

Unstructured **G Matrix**

(RANDOM statement)

Each person has same **1 x 1 G** matrix (no covariance across persons in two-level model)

Random Intercept Variance only

$$\begin{bmatrix} \tau^2_{U_0} \end{bmatrix}$$

Level 1, WP Variance

Diagonal (VC) **R Matrix**

(REPEATED statement)

Each person has same ***n* x *n* R** matrix → **equal variances and 0 covariances** across time (no covariance across persons)

Residual Variance only

$$\begin{bmatrix} \sigma^2_e & 0 & 0 & 0 \\ 0 & \sigma^2_e & 0 & 0 \\ 0 & 0 & \sigma^2_e & 0 \\ 0 & 0 & 0 & \sigma^2_e \end{bmatrix}$$

Building V: Fixed Linear Time, Random Intercept Model (*n* = 4 occasions)

Scalar "mixed" model equation per person:

$$\mathbf{Y}_i = \boxed{\mathbf{X}_i * \boldsymbol{\gamma}} + \boxed{\mathbf{Z}_i * \mathbf{U}_i + \mathbf{E}_i}$$

$$\begin{pmatrix} y_{0i} \\ y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \gamma_{00} \\ \gamma_{10} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} U_{0i} \end{pmatrix} + \begin{pmatrix} e_{0i} \\ e_{1i} \\ e_{2i} \\ e_{3i} \end{pmatrix}$$

$$\begin{pmatrix} y_{0i} \\ y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix} = \begin{pmatrix} \gamma_{00} + \gamma_{10}(0) \\ \gamma_{00} + \gamma_{10}(1) \\ \gamma_{00} + \gamma_{10}(2) \\ \gamma_{00} + \gamma_{10}(3) \end{pmatrix} + \begin{pmatrix} U_{0i} \\ U_{0i} \\ U_{0i} \\ U_{0i} \end{pmatrix} + \begin{pmatrix} e_{0i} \\ e_{1i} \\ e_{2i} \\ e_{3i} \end{pmatrix}$$

$$\begin{pmatrix} y_{0i} \\ y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix} = \begin{pmatrix} \gamma_{00} + \gamma_{10}(0) + U_{0i} + e_{0i} \\ \gamma_{00} + \gamma_{10}(1) + U_{0i} + e_{1i} \\ \gamma_{00} + \gamma_{10}(2) + U_{0i} + e_{2i} \\ \gamma_{00} + \gamma_{10}(3) + U_{0i} + e_{3i} \end{pmatrix}$$

$\mathbf{X}_i$ = *n* x *k* values of **predictors with fixed effects**, so values differ per person (*k* = 2: intercept, linear time)

$\boldsymbol{\gamma}$ = *k* x 1 estimated **fixed effects**, so will be the same for everyone (γ_{00} = intercept, γ_{10} = linear time)

$\mathbf{Z}_i$ = *n* x *u* values of **predictors with random effects**, so values differ per person (*u* = 1: intercept)

$\mathbf{U}_i$ = *u* x 1 estimated **random effects**, so values differ per person

$\mathbf{E}_i$ = *n* x *n* **time-specific residuals**, so values differ per person

Building V: Fixed Linear Time, Random Intercept Model (n = 4 occasions)

Predicted variances and covariances per person:

$$\boxed{\mathbf{V}_i = \mathbf{Z}_i * \mathbf{G}_i * \mathbf{Z}_i^T + \mathbf{R}_i}$$

$$\mathbf{V}_i = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}\left(\tau^2_{U_0}\right)\begin{pmatrix}1 & 1 & 1 & 1\end{pmatrix} + \begin{pmatrix}\sigma^2_e & 0 & 0 & 0\\ 0 & \sigma^2_e & 0 & 0\\ 0 & 0 & \sigma^2_e & 0\\ 0 & 0 & 0 & \sigma^2_e\end{pmatrix}$$

$\mathbf{V}_i$: Variance$\left[y_{time}\right]$

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

$\mathbf{V}_i$: Covariance$\left[y_A, y_B\right]$

$= \tau^2_{U_0}$

$$\mathbf{V}_i = \begin{bmatrix}\sigma^2_e+\tau^2_{u_0} & \tau^2_{u_0} & \tau^2_{u_0} & \tau^2_{u_0}\\ \tau^2_{u_0} & \sigma^2_e+\tau^2_{u_0} & \tau^2_{u_0} & \tau^2_{u_0}\\ \tau^2_{u_0} & \tau^2_{u_0} & \sigma^2_e+\tau^2_{u_0} & \tau^2_{u_0}\\ \tau^2_{u_0} & \tau^2_{u_0} & \tau^2_{u_0} & \sigma^2_e+\tau^2_{u_0}\end{bmatrix}$$

$\mathbf{Z}_i$ = n x u values of **predictors with random effects**, so values differ per person (u = 1: intercept)

$\mathbf{Z}_i^T$ = u x n values of predictors with random effects (just $\mathbf{Z}_i$ transposed)

$\mathbf{G}_i$ = u x u estimated **random effects variances and covariances**, so will be the same for everyone (τ_{U0}^2 = intercept variance)

$\mathbf{R}_i$ = n x n **time-specific residual variances and covariances**, so will be same for everyone (diagonal here)

Fixed Linear, Random Linear Time Model:

$$y_{ti} = \gamma_{00} + U_{0i} + (\gamma_{10}\text{Time}_{ti}) + \boxed{(U_{1i}\text{Time}_{ti})} + e_{ti}$$

Total Predicted Data Matrix is called V Matrix

$$\begin{bmatrix}\hat{\sigma}^2_{11} & \hat{\sigma}_{12} & \hat{\sigma}_{13} & \hat{\sigma}_{14}\\ \hat{\sigma}_{21} & \hat{\sigma}^2_{22} & \hat{\sigma}_{23} & \hat{\sigma}_{24}\\ \hat{\sigma}_{31} & \hat{\sigma}_{32} & \hat{\sigma}^2_{33} & \hat{\sigma}_{43}\\ \hat{\sigma}_{41} & \hat{\sigma}_{42} & \hat{\sigma}_{43} & \hat{\sigma}^2_{44}\end{bmatrix}$$

n = 4 occasions

Level 2, BP Variance

Unstructured **G Matrix**
(RANDOM statement)
Each person has same **2 x 2 G** matrix (no covariance across persons in two-level model)

Random Intercept Variance

Random Slope Variance

$$\begin{bmatrix}\tau^2_{U_0} & \tau_{U_{10}}\\ \tau_{U_{01}} & \tau^2_{U_1}\end{bmatrix}$$

Level 1, WP Variance

Diagonal (VC) **R Matrix**
(REPEATED statement)
Each person has same ***n* x *n* R** matrix → **equal variances and 0 covariances** across time
(no covariance across persons)

Residual Variance only

$$\begin{bmatrix}\sigma^2_e & 0 & 0 & 0\\ 0 & \sigma^2_e & 0 & 0\\ 0 & 0 & \sigma^2_e & 0\\ 0 & 0 & 0 & \sigma^2_e\end{bmatrix}$$

Building **V**: **Fixed Linear**, **Random Linear** Time Model (n = 4)

Scalar "mixed" model equation per person:

$$\mathbf{Y}_i = \boxed{\mathbf{X}_i * \boldsymbol{\gamma}} + \boxed{\mathbf{Z}_i * \mathbf{U}_i + \mathbf{E}_i}$$

$$\begin{pmatrix} y_{0i} \\ y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \gamma_{00} \\ \gamma_{10} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} U_{0i} \\ U_{1i} \end{pmatrix} + \begin{pmatrix} e_{0i} \\ e_{1i} \\ e_{2i} \\ e_{3i} \end{pmatrix}$$

$$\begin{pmatrix} y_{0i} \\ y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix} = \begin{pmatrix} \gamma_{00} + \gamma_{10}(0) \\ \gamma_{00} + \gamma_{10}(1) \\ \gamma_{00} + \gamma_{10}(2) \\ \gamma_{00} + \gamma_{10}(3) \end{pmatrix} + \begin{pmatrix} U_{0i} + U_{1i}(0) \\ U_{0i} + U_{1i}(1) \\ U_{0i} + U_{1i}(2) \\ U_{0i} + U_{1i}(3) \end{pmatrix} + \begin{pmatrix} e_{0i} \\ e_{1i} \\ e_{2i} \\ e_{3i} \end{pmatrix}$$

$$\begin{pmatrix} y_{0i} \\ y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix} = \begin{pmatrix} \gamma_{00} + \gamma_{10}(0) + U_{0i} + U_{1i}(0) + e_{0i} \\ \gamma_{00} + \gamma_{10}(1) + U_{0i} + U_{1i}(1) + e_{1i} \\ \gamma_{00} + \gamma_{10}(2) + U_{0i} + U_{1i}(2) + e_{2i} \\ \gamma_{00} + \gamma_{10}(3) + U_{0i} + U_{1i}(3) + e_{3i} \end{pmatrix}$$

$\mathbf{X}_i$ = n x k values of **predictors with fixed effects**, so values differ per person (k = 2: intercept, linear time)

$\boldsymbol{\gamma}$ = k x 1 estimated **fixed effects**, so will be the same for everyone (γ_{00} = intercept, γ_{10} = linear time)

$\mathbf{Z}_i$ = n x u values of **predictors with random effects**, so values differ per person (u = 2: intercept, linear time)

$\mathbf{U}_i$ = u x 1 estimated **random effects**, so values differ per person

$\mathbf{E}_i$ = n x n **time-specific residuals**, so values differ per person

Building **V**: Fixed Linear, **Random Linear** Time Model (n = 4)

Predicted variances and covariances per person:

$$\mathbf{V}_i = \boxed{\mathbf{Z}_i * \mathbf{G}_i * \mathbf{Z}_i^T + \mathbf{R}_i}$$

$$\mathbf{V}_i = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \tau^2_{U_0} & \tau_{U_{01}} \\ \tau_{U_{01}} & \tau^2_{U_1} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} + \begin{pmatrix} \sigma^2_e & 0 & 0 & 0 \\ 0 & \sigma^2_e & 0 & 0 \\ 0 & 0 & \sigma^2_e & 0 \\ 0 & 0 & 0 & \sigma^2_e \end{pmatrix}$$

$\mathbf{V}_i$ matrix = complicated ☺

$\mathbf{V}_i$ matrix: Variance$[y_{time}]$

$$= \tau^2_{U_0} + \left[(\text{time})^2 \tau^2_{U_1}\right] + \left[2(\text{time})\tau_{U_{01}}\right] + \sigma^2_e$$

$\mathbf{V}_i$ matrix: Covariance$[y_A, y_B]$

$$= \tau^2_{U_0} + \left[(A+B)\tau_{U_{01}}\right] + \left[(AB)\tau^2_{U_1}\right]$$

$\mathbf{Z}_i$ = n x u values of **predictors with random effects**, so values differ per person (u = 2: int., time slope)

$\mathbf{Z}_i^T$ = u x n values of predictors with random effects (just $\mathbf{Z}_i$ transposed)

$\mathbf{G}_i$ = u x u estimated **random effects variances and covariances**, so will be the same for everyone (τ_{U0}^2 = int. variance, τ_{U1}^2 = linear slope variance)

$\mathbf{R}_i$ = n x n **time-specific residual variances and covariances**, so will be same for everyone (diagonal here)

V Across Persons: **Random Linear** Time Model (*n* = 4)

- **V** for two persons with unbalanced time observations:

$$\mathbf{V} = \mathbf{Z} * \mathbf{G} * \mathbf{Z}^T + \mathbf{R}$$

$$\mathbf{V} = \begin{pmatrix} 1 & 0.0 & 0 & 0 \\ 1 & 1.0 & 0 & 0 \\ 1 & 2.0 & 0 & 0 \\ 1 & 3.0 & 0 & 0 \\ 0 & 0 & 1 & 0.2 \\ 0 & 0 & 1 & 1.4 \\ 0 & 0 & 1 & 2.3 \\ 0 & 0 & 1 & 3.5 \end{pmatrix} \begin{pmatrix} \tau^2_{U_0} & \tau_{U_{01}} & 0 & 0 \\ \tau_{U_{01}} & \tau^2_{U_1} & 0 & 0 \\ 0 & 0 & \tau^2_{U_0} & \tau_{U_{01}} \\ 0 & 0 & \tau_{U_{01}} & \tau^2_{U_1} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0.0 & 1.0 & 2.0 & 3.0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0.2 & 1.4 & 2.3 & 3.5 \end{pmatrix} + \begin{pmatrix} \sigma^2_e & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2_e & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2_e & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2_e & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2_e & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma^2_e & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_e & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_e \end{pmatrix}$$

- The giant combined **V** matrix across persons is how the multilevel or mixed model is actually estimated
  - Known as "block diagonal" structure → stuff for each person, but 0's for the elements that describe relationships between persons (because persons are supposed to be independent here!)
    - The "block diagonal" does not need to be the same size or contain the same time observations per person…

The Take-Home Point

- The partitioning of variance into piles…
  - Level 2 = BP → **G** matrix of random effects variances/covariances
  - Level 1 = WP → **R** matrix of residual variances/covariances
  - **G** and **R** combine to create **V** matrix of total variances/covariances
  - Many flexible options that allows the variances and covariances to vary in a time-dependent way that better matches the actual data

Compound Symmetry (CS)

$$\begin{bmatrix} \sigma^2_e+\tau^2_{u_0} & \tau^2_{u_0} & \tau^2_{u_0} & \tau^2_{u_0} \\ \tau^2_{u_0} & \sigma^2_e+\tau^2_{u_0} & \tau^2_{u_0} & \tau^2_{u_0} \\ \tau^2_{u_0} & \tau^2_{u_0} & \sigma^2_e+\tau^2_{u_0} & \tau^2_{u_0} \\ \tau^2_{u_0} & \tau^2_{u_0} & \tau^2_{u_0} & \sigma^2_e+\tau^2_{u_0} \end{bmatrix}$$

MLM uses **G** and **R** to predict something in-between!

Unstructured (UN)

$$\begin{bmatrix} \sigma^2_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma^2_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma^2_{33} & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma^2_{44} \end{bmatrix}$$

Parsimony ←→ Good fit

Univariate RM ANOVA

NAME …THAT … STRUCTURE!

Multivariate RM ANOVA

Summary: The Point of MLM

- All we've done so far is **carve up** our total variance into up to 3 piles:
  - BP (error) variance around intercept
  - BP (error) variance around slope
  - WP (error) residual variance

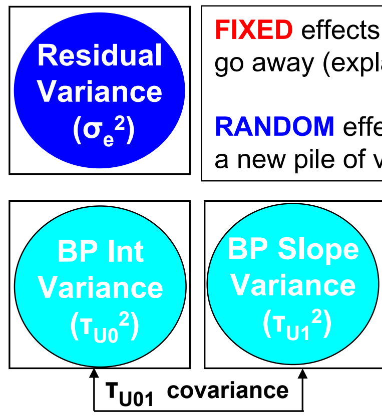
  These two are one pile of "error variance" in RM Anova
- **But making piles does not make error variance go away…**

Level 1 (one source of) Within-Person Variation:

Gets accounted for by time-level predictors

Level 2 (two sources of) Between-Person Variation:

Gets accounted for by person-level predictors



FIXED effects make variance go away (explain variance).

RANDOM effects just make a new pile of variance.

Lecture 2: Concepts in Modeling Within-Person Change and Fluctuation

- Empty Models and Intraclass Correlation
- Fixed vs. Random Effects of Persons and Time
- **Considering Alternative Metrics of Time**
- Within-Person Change vs. Within-Person Fluctuation

Modeling Within-Person Change

- The goal of creating statistical models of change is to describe the overall pattern of and predict individual differences in **change over time**.

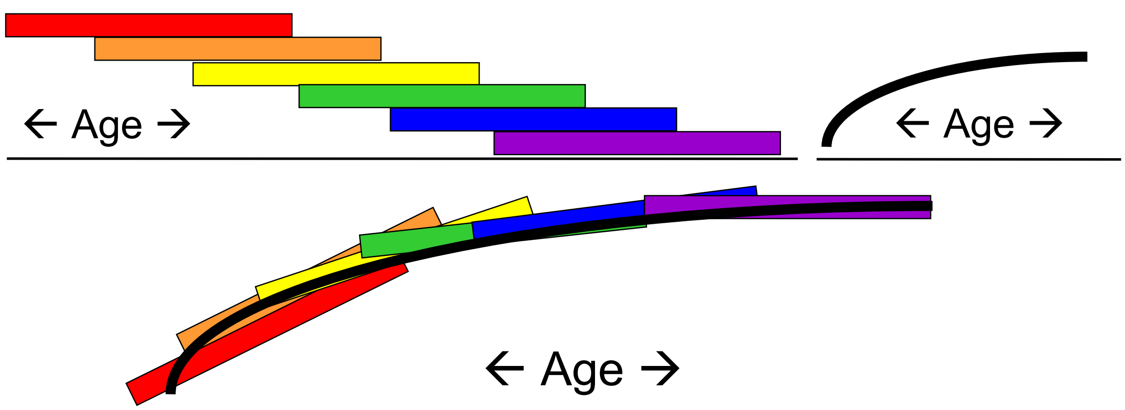
- Longitudinal models employ an often unrecognized assumption that we know exactly **what "time" should be**.

- This involves 2 related concerns:
  - **What should "time" be?**
  - **What do we do when people differ in "time"?**
    - Between-person effects vs. within-person effects of time
  - Concerns apply to *accelerated longitudinal designs*

Accelerated Longitudinal Designs

Want to do a longitudinal study but just don't have the time?

Accelerate: Model trajectories over a wider span of time than directly possible using only observed longitudinal information



Does anybody really know what *Time* is*?

- **First: What should “time” be?**
  - What is the **causal process** by which we are indexing change?
  - What do we do when **multiple processes** may be at work?
  - Relevant for merging different persons onto same time metric, but not a relevant distinction within-persons

- Consider some examples…
  - Growth in scholastic achievement in children
  - Improvement in job performance of employees
  - Changes in marital satisfaction in spouses
  - Physical and cognitive decline in older adults

** Title with thanks to Chicago*

What is ‘time’? *Occasion of Study*

- What I tend to use most of the time…
- Can be used generically for many purposes
- Represents within-person changes directly
- Makes no assumptions about what underlying process might be responsible for observed change

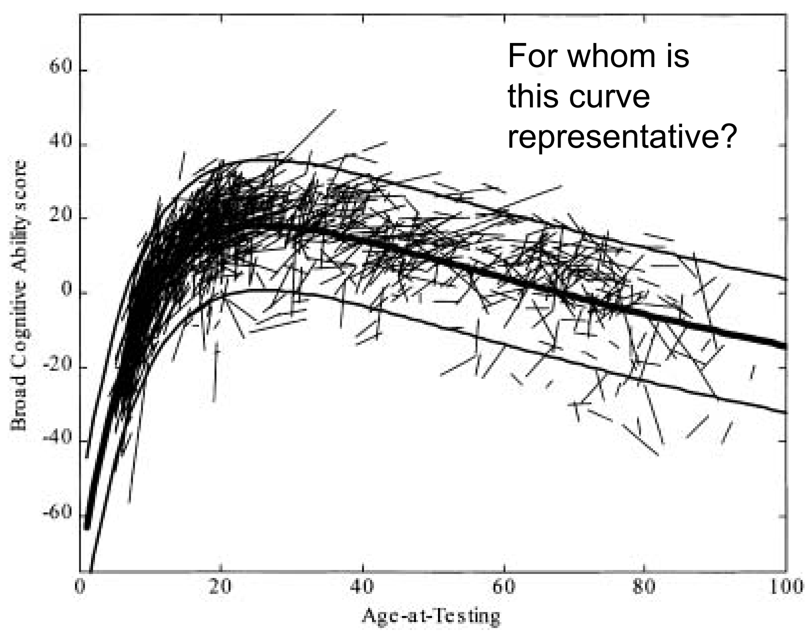
- Differences in age at baseline (or in other between-person time-related variables) can then be used as <u>predictors</u> of change rather than as the <u>metric</u> of time:
  - Older individuals may be higher/lower initially
  - Older individuals may change more/less steeply over time
  - These are testable hypotheses!

What is ‘time’? *Age-as-Time*

- If people differ in initial age, measuring change as a function of age requires assuming **age convergence**
  - Younger people and older people differ *only* by age
  - Between-person, cross-sectional *age effects* are equivalent to within-person, longitudinal *aging effects*
- Age convergence is not likely to hold when
  - Initial age range is large
  - Cohort differences and selection effects are large
- Age convergence can and should be tested empirically!
  - Applies to any time metric with both BP and WP variation

Broad Cognitive Ability Longitudinal Data and Expectations

(McArdle, Ferrar-Caja, Hamagami, & Woodcock, 2002)

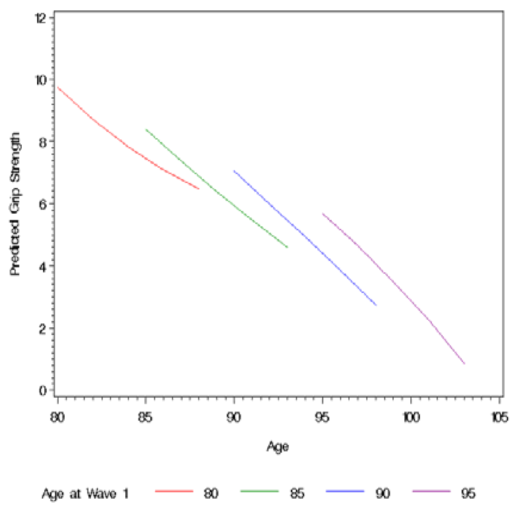


This model combines cross-sectional and longitudinal information...

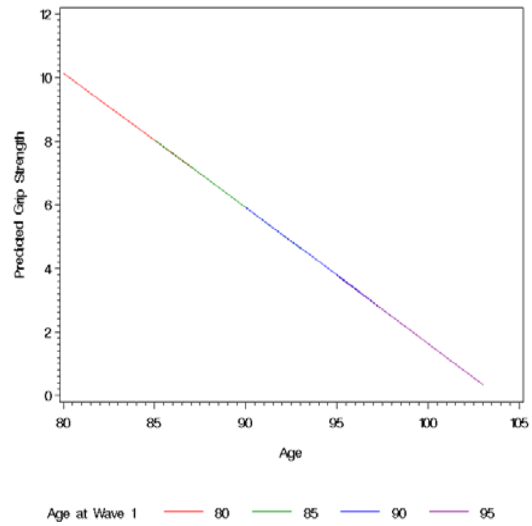
But what is the **relative contribution** of cross-sectional to longitudinal information?

Grip Strength in the Oldest-Old

Time in Study: Age as Predictor



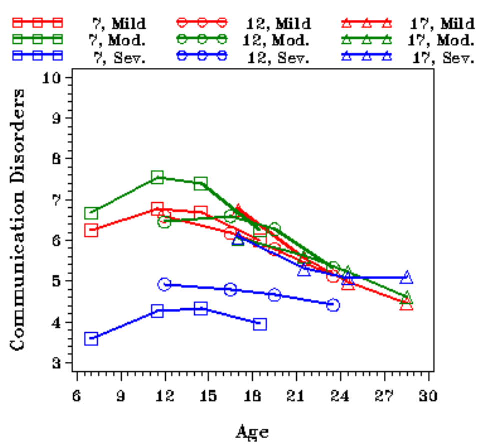
Age as Time



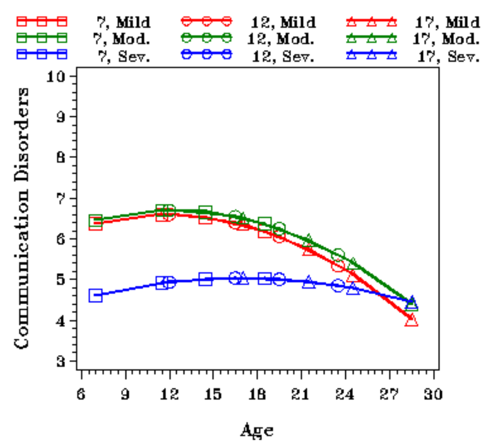
When age convergence does not obtain, age-based models can be misleading because you have squished distinct trajectories together artificially **(cross-sectional ≠ longitudinal)**.

Communications Disturbance by Age (4-19 at entry) and ID

Time-In-Study Trajectories



Age-as-Time Trajectories

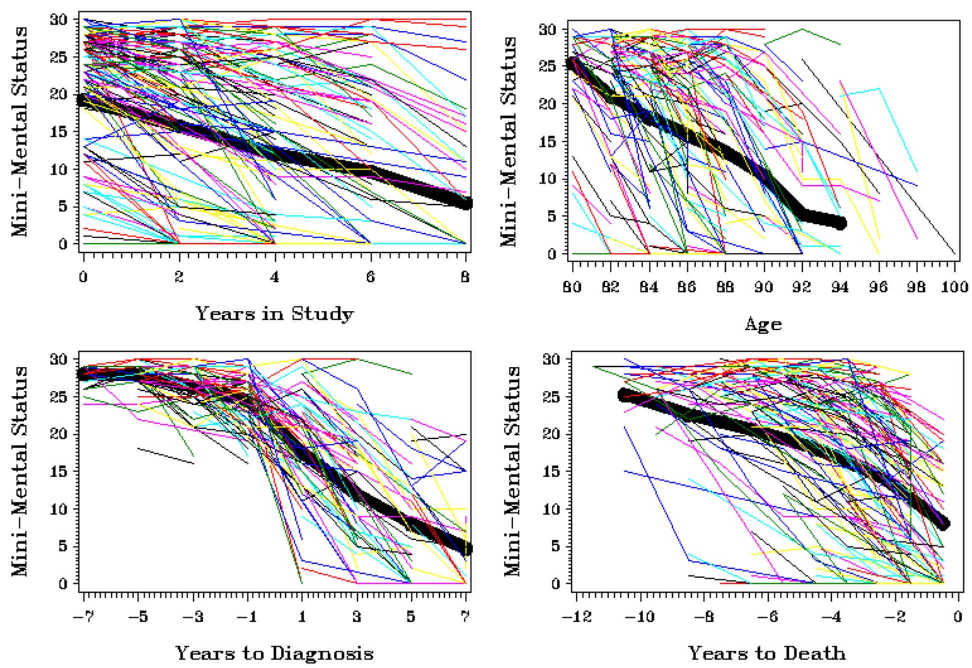


<u>Non-convergence:</u> People from different age cohorts are not on the same trajectories, and do not have the same expected values at the same age. **Cross-sectional ≠ Longitudinal!**

What is ‘time’? *Distance to/from a Common Important Event*

- Useful if change is due to an ongoing process
  - Time since Diagnosis, Time to Death, Time since Divorce, Time since GED, Time in Therapy
- Different ‘Times’ can contribute in different ways:
  - Age-as-Time, but with Initial Diagnosis Time as a predictor
    - Initial status and rate of change are driven by *age*, but vary their effects vary by *amount of disease progression*
  - Diagnosis-as-Time, but with Initial Age as a predictor
    - Initial status and rate of change are driven by *amount of disease progression*, but vary by initial age

Individual Trajectories over “Time”



Lecture 2: Concepts in Modeling Within-Person Change and Fluctuation

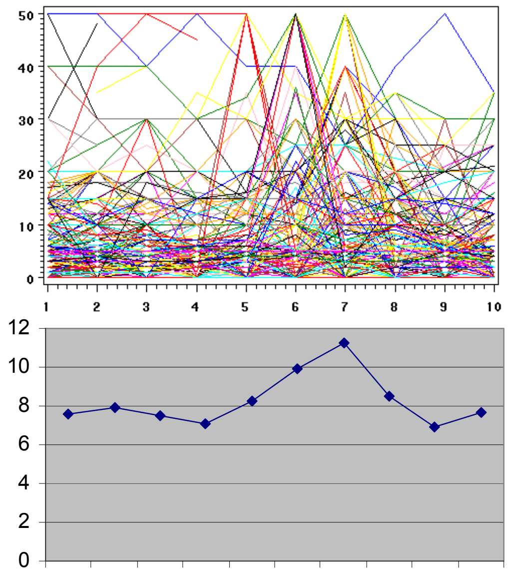
- Empty Models and Intraclass Correlation
- Fixed vs. Random Effects of Persons and Time
- Considering Alternative Metrics of Time
- **Within-Person Change vs. Within-Person Fluctuation**

Does anybody really care (about *Time*)?

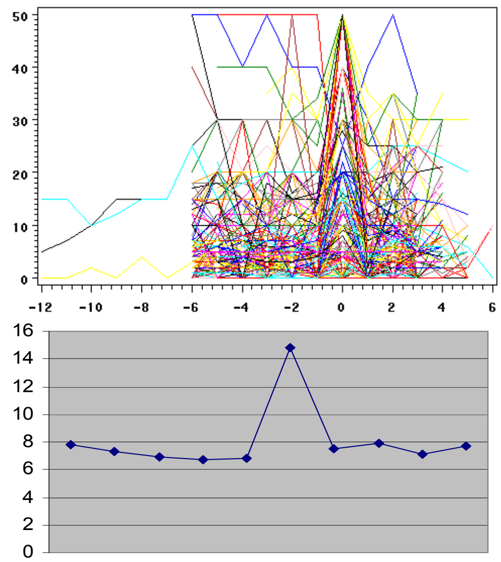
- Even in longitudinal studies focused on **within-person fluctuation rather than change**, time may still be relevant
- For instance, in daily diary studies:
  - Day of the Week (time metric could be **day of week**)
  - Fatigue/Reactivity (time metric could be **day of study**)
- In these cases you'd be "controlling for change" instead of "modeling change" (same models, different emphasis)
  - Some examples…

Plans to Drink Alcohol: Time-in-Study vs. Time-to-Event

#Drinks by Interview Week Number

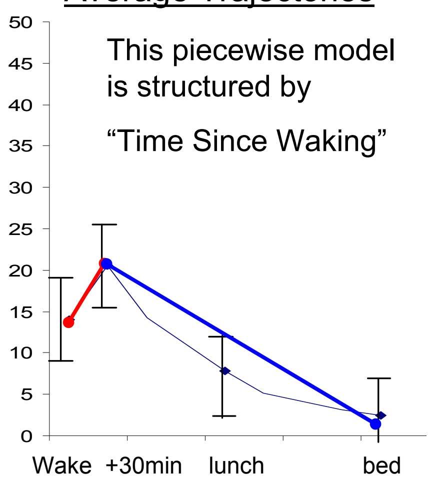


#Drinks by Time to Spring Break



Daily Fluctuation in Cortisol: Morning Rise and Afternoon Fall

<u>Average Trajectories</u>



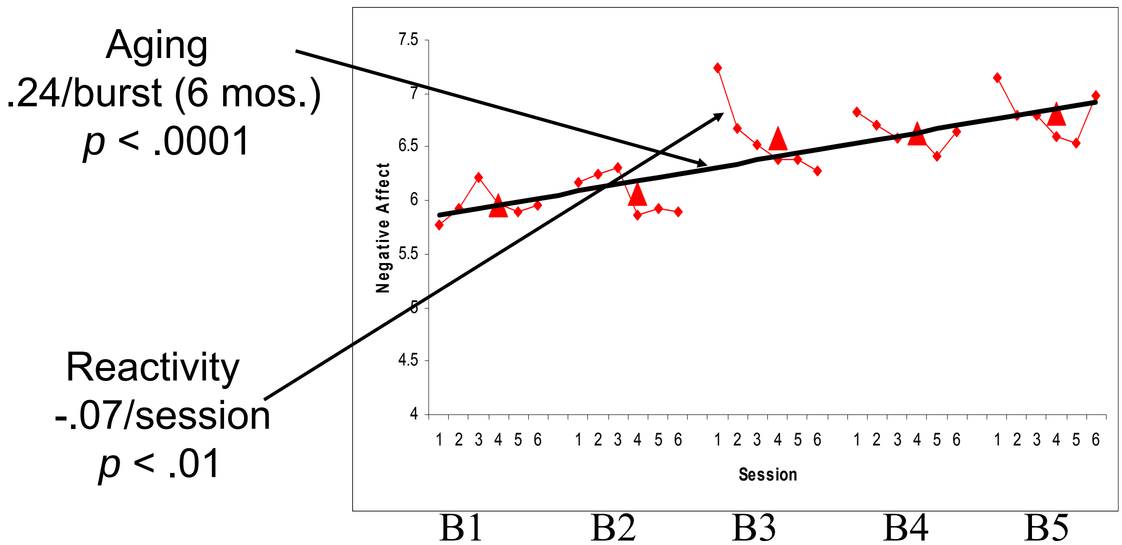
<u>SAS Code to create two piecewise slopes from continuous time of day in stacked data:</u>

```
IF obs=1 THEN DO;
    P1=0;              P2=0; END;
IF obs=2 THEN DO;
    P1= time2-time1; P2=0; END;
IF obs=3 THEN DO;
    P1= time2-time1; P2=time3-time2; END;
IF obs=4 THEN DO;
    P1= time2-time1; P2=time4-time2; END;
```

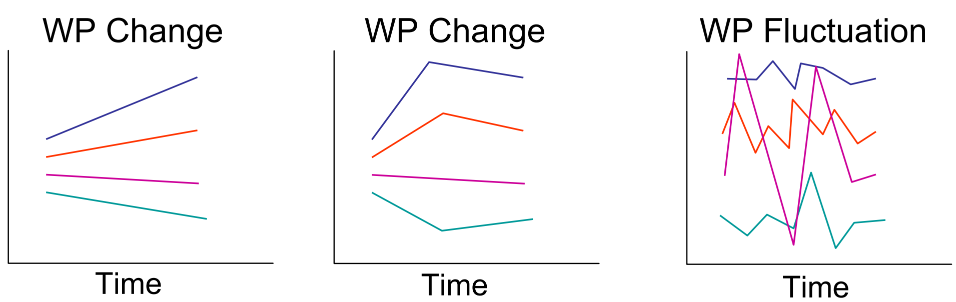
Waking time could be a level-2 predictor!

Change in Negative Affect over "Time"

Stawski & Sliwinski, GSA 2005



Modeling Within-Person Change vs. Within-Person Fluctuation



Model for the Means:

- WP Change → describe pattern of *average* change (over "time")
- WP Fluctuation → *may* not need anything (if no systematic change)

Model for the Variances:

- WP Change → describe *individual differences* in change (random effects)
  → this allows variances and covariances to differ over time
- WP Fluctuation → describe pattern of variances and covariances over time