PSY792F SEM

Week 12 – Multilevel Modeling

Lara N. Pantlin, MS

Mark A. Prince, PhD, MS

Multilevel Modeling (MLM)

- Different names:
    - Multilevel regression models
    - Random coefficient models
    - Variance component models
    - Latent growth models
    - Mixed linear models
- MLM is used when there is **more than one level of nesting** in your data
    - Students within classrooms
    - Assessment of individuals across time
    - Patients assigned to doctors (Hayes, 2006)

Why do you need to use MLM?

- Linear regression approaches assume:
  - Independent observations
  - Independent error terms
  - Equal variances of errors for all observations
- With hierarchical (i.e., nested) data
  - Observation are not independent
    - People in a group tend to be more like others in their group than they are like others in a different group.
    - Selection, shared history, shared experiences, common geography
  - Errors are not independent
  - Units in different clusters may have different variances

MLM vs. RM-ANOVA

1. MLM has Less Stringent Assumptions: MLM can be used if the assumptions of constant variances (homogeneity of variance, or homoscedasticity), constant covariances (compound symmetry), or constant variances of differences scores (sphericity) are violated for RM-ANOVA.

2. MLM Allows Hierarchical Structure: MLM can be used for higher-order sampling procedures, whereas RM-ANOVA is limited to examining two-level sampling procedures. In other words, MLM can look at repeated measures within subjects, within a third level of analysis etc., whereas RM-ANOVA is limited to repeated measures within subjects.

MLM vs. RM-ANOVA Continued

3. MLM can Handle Missing Data: Missing data is permitted in MLM without causing additional complications. With RM-ANOVA, subject's data must be excluded if they are missing a single data point. Missing data and attempts to resolve missing data (i.e. using the subject's mean for non-missing data) can raise additional problems in RM-ANOVA. *Note:* Although missing data is permitted in MLM, it is assumed to be missing at random. Thus, systematically missing data can present problems.

4. MLM can also handle data in which there is variation in the exact timing of data collection (i.e. variable timing versus fixed timing). For example, data for a longitudinal study may attempt to collect measurements at age 6 months, 9 months, 12 months, and 15 months. However, participant availability, bank holidays, and other scheduling issues may result in variation regarding when data is collected. This variation may be addressed in MLM by adding time into the regression equation. There is also no need for equal intervals between measurement points in MLM.

5. MLM is relatively easily extended to discrete (categorical) data.

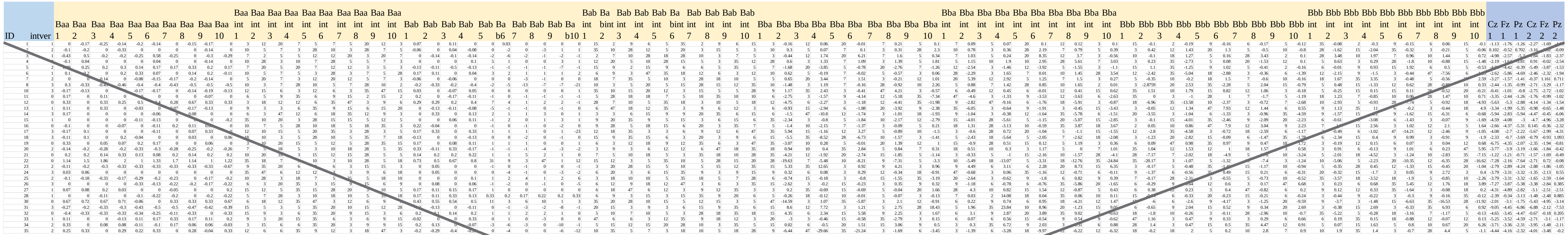
Data structure

- Data need to be in long/tall format
  - Each participant should have multiple rows of data
  - There should be an indicator for time or observation number

Wide Data Format				
ID	Alc1	Alc2	Alc3	Sex
1	5	5	5	0
2	0	0	8	1
3	1	2	1	0

Tall Data Format			
ID	Alc	time	sex
1	5	1	0
1	5	2	0
1	5	3	0
2	0	1	1
2	0	2	1
2	8	3	1
3	1	1	0
3	2	2	0
3	1	3	0

Wide Format: For reference file name: WidetoLongIntervalDemo Headers.xlsx



ID	Inter	Baa1	…	Baa10	Baa Int1	…	Baa Int10	Bab1	…	Bab10	Bab Int1	…	Bab Int10	Bba1	…	Bba 10	…	Bba Int10	Bbb1	…	Bbb 10	Bbb Int1	…	Bbb Int10	Cz1	Fz1	Pz1	Cz2	Fz2	Pz2

```
DATA WIDETOLONG:
   WIDE = Baa1-Baa10 | BaaInt1-BaaInt10 | Bab1-Bab10 |
   BabInt1-BabInt10 | Bba1-Bba10 |
   BbaInt1-BbaInt10 | Bbb1-Bbb10 | BbbInt1-BbbInt10;
LONG = BaaT1 | BaaE1I | BabT2 | BabE2I | BbaT1 | BbaR1I | BbbT2 | BbbR2I;
   IDVARIABLE = ID;
   REPETITION = trial;
   SAVEDATA: FILE IS Demolong.dat;
```

```
DATA: FILE IS WidetoLongIntervalDemo.csv;

DATA WIDETOLONG:
   WIDE = Baa1-Baa10 | BaaInt1-BaaInt10 | Bab1-Bab10 |
   BabInt1-BabInt10 | Bba1-Bba10 |
   BbaInt1-BbaInt10 | Bbb1-Bbb10 | BbbInt1-BbbInt10;
LONG = BaaT1 | BaaE1I | BabT2 | BabE2I | BbaT1 | BbaR1I | BbbT2 |
BbbR2I;
   IDVARIABLE = ID;
   REPETITION = trial;
   SAVEDATA: FILE IS Demolong.dat;

VARIABLE:
NAMES ARE  ID interven Baa1-Baa10 BaaInt1-BaaInt10 Bab1-Bab10
BabInt1-BabInt10 Bba1-Bba10 BbaInt1-BbaInt10 Bbb1-Bbb10
BbbInt1-BbbInt10 cz1 fz1 pz1 cz2 fz2 pz2;

USEVARIABLES ARE  ID interven cz1 fz1 pz1 cz2 fz2 pz2
BaaT1 BaaE1I BabT2 BabE2I BbaT1 BbaR1I BbbT2 BbbR2I trial;

CLUSTER = ID;

ANALYSIS: TYPE is twolevel basic;
```

When in wide format, 1 participant has 10 trials.

```
DATA: FILE IS WidetoLongIntervalDemo.csv;

DATA WIDETOLONG:
    WIDE = Baa1-Baa10 | BaaInt1-BaaInt10 | Bab1-Bab10 |
    BabInt1-BabInt10 | Bba1-Bba10 |
    BbaInt1-BbaInt10 | Bbb1-Bbb10 | BbbInt1-BbbInt10;
LONG = BaaT1 | BaaE1I | BabT2 | BabE2I | BbaT1 | BbaR1I | BbbT2
| BbbR2I;
    IDVARIABLE = ID;
    REPETITION = trial;
    SAVEDATA: FILE IS Demolong.dat;

VARIABLE:
NAMES ARE  ID interven Baa1-Baa10 BaaInt1-BaaInt10 Bab1-Bab10
BabInt1-BabInt10 Bba1-Bba10 BbaInt1-BbaInt10 Bbb1-Bbb10
BbbInt1-BbbInt10 cz1 fz1 pz1 cz2 fz2 pz2;

USEVARIABLES ARE  ID interven cz1 fz1 pz1 cz2 fz2 pz2
BaaT1 BaaE1I BabT2 BabE2I BbaT1 BbaR1I BbbT2 BbbR2I trial;

CLUSTER = ID;

ANALYSIS: TYPE is twolevel basic;
```

Repeated in NAMES ARE These variables will change i.e. 10 trials with different values

```
DATA: FILE IS WidetoLongIntervalDemo.csv;

DATA WIDETOLONG:
    WIDE = Baa1-Baa10 | BaaInt1-BaaInt10 | Bab1-Bab10 |
    BabInt1-BabInt10 | Bba1-Bba10 |
    BbaInt1-BbaInt10 | Bbb1-Bbb10 | BbbInt1-BbbInt10;
LONG = BaaT1 | BaaE1I | BabT2 | BabE2I | BbaT1 | BbaR1I | BbbT2 | BbbR2I;
    IDVARIABLE = ID;
    REPETITION = trial;
    SAVEDATA: FILE IS Demolong.dat;

VARIABLE:
NAMES ARE  ID interven Baa1-Baa10 BaaInt1-BaaInt10 Bab1-Bab10
BabInt1-BabInt10 Bba1-Bba10 BbaInt1-BbaInt10 Bbb1-Bbb10
BbbInt1-BbbInt10 cz1 fz1 pz1 cz2 fz2 pz2;

USEVARIABLES ARE  ID interven cz1 fz1 pz1 cz2 fz2 pz2
BaaT1 BaaE1I BabT2 BabE2I BbaT1 BbaR1I BbbT2 BbbR2I trial;

CLUSTER = ID;

ANALYSIS: TYPE is twolevel basic;
```

These variables will not change. The same number will be repeated 10 times.

```
DATA: FILE IS WidetoLongIntervalDemo.csv;

DATA WIDETOLONG:
    WIDE = Baa1-Baa10 | BaaInt1-BaaInt10 | Bab1-Bab10 |
    BabInt1-BabInt10 | Bba1-Bba10 |
    BbaInt1-BbaInt10 | Bbb1-Bbb10 | BbbInt1-BbbInt10;
LONG = BaaT1 | BaaE1I | BabT2 | BabE2I | BbaT1 | BbaR1I | BbbT2 | BbbR2I;
    IDVARIABLE = ID;
    REPETITION = trial;
    SAVEDATA: FILE IS Demolong.dat;

VARIABLE:
NAMES ARE  ID interven Baa1-Baa10 BaaInt1-BaaInt10 Bab1-Bab10
BabInt1-BabInt10 Bba1-Bba10 BbaInt1-BbaInt10 Bbb1-Bbb10
BbbInt1-BbbInt10 cz1 fz1 pz1 cz2 fz2 pz2;

USEVARIABLES ARE  ID interven cz1 fz1 pz1 cz2 fz2 pz2
BaaT1 BaaE1I BabT2 BabE2I BbaT1 BbaR1I BbbT2 BbbR2I trial;

CLUSTER = ID;

ANALYSIS: TYPE is twolevel basic;
```

Comes first in USEVARIABLES

```
DATA: FILE IS WidetoLongIntervalDemo.csv;

DATA WIDETOLONG:
   WIDE = Baa1-Baa10 | BaaInt1-BaaInt10 | Bab1-Bab10 |
   BabInt1-BabInt10 | Bba1-Bba10 |
   BbaInt1-BbaInt10 | Bbb1-Bbb10 | BbbInt1-BbbInt10;
LONG = BaaT1 | BaaE1I | BabT2 | BabE2I | BbaT1 | BbaR1I |
BbbT2 | BbbR2I;
   IDVARIABLE = ID;
   REPETITION = trial;
   SAVEDATA: FILE IS Demolong.dat;

VARIABLE:
NAMES ARE  ID interven Baa1-Baa10 BaaInt1-BaaInt10 Bab1-
Bab10
BabInt1-BabInt10 Bba1-Bba10 BbaInt1-BbaInt10 Bbb1-Bbb10
BbbInt1-BbbInt10 cz1 fz1 pz1 cz2 fz2 pz2;

USEVARIABLES ARE  ID interven cz1 fz1 pz1 cz2 fz2 pz2
BaaT1 BaaE1I BabT2 BabE2I BbaT1 BbaR1I BbbT2 BbbR2I trial;

CLUSTER = ID;

ANALYSIS: TYPE is twolevel basic;
```

WIDE:

ID	intver	Baa 1	Baa 2	Baa 3	Baa 4	Baa 5	Baa 6	Baa 7	Baa 8	Baa 9	Baa 10
1	1	0	-0.17	-0.25	-0.14	-0.2	-0.14	0	-0.15	-0.17	0
2	2	-0.1	-0.2	0	-0.33	0	0	0	0	-0.14	0
3	1	-0.43	0.5	-0.2	-0.2	-0.25	0.58	-0.25	0	-0.3	-0.29
4	2	-0.1	0.04	0	0	0	0.04	0	0	-0.14	0
5	2	0.29	0.25	0.2	0.3	0.14	0.17	0.17	0.33	0.2	0.17
6	1	0.1	0.2	0	0.2	0.33	0.07	0	0.14	0.2	-0.11
7	2	0	0	-0.14	0	-0.08	-0.15	-0.17	-0.2	-0.14	0
8	3	0.3	-0.33	-0.43	-0.46	-0.4	-0.4	-0.43	-0.5	-0.5	-0.5

LONG:

ID	Inter	BaaT1
1	1	0
1	1	-0.17
1	1	-0.25
1	1	-0.14
1	1	-0.2
1	1	-0.14
1	1	0
1	1	-0.15
1	1	-0.17
1	1	0

```
DATA: FILE IS WidetoLongIntervalDemo.csv;

DATA WIDETOLONG:
   WIDE = Baa1-Baa10 | BaaInt1-BaaInt10 | Bab1-Bab10 |
   BabInt1-BabInt10 | Bba1-Bba10 |
   BbaInt1-BbaInt10 | Bbb1-Bbb10 | BbbInt1-BbbInt10;
LONG = BaaT1 | BaaE1I | BabT2 | BabE2I | BbaT1 | BbaR1I | BbbT2 |
BbbR2I;
   IDVARIABLE = ID;
   REPETITION = trial;
   SAVEDATA: FILE IS Demolong.dat;

VARIABLE:
NAMES ARE  ID interven Baa1-Baa10 BaaInt1-BaaInt10 Bab1-Bab10
BabInt1-BabInt10 Bba1-Bba10 BbaInt1-BbaInt10 Bbb1-Bbb10
BbbInt1-BbbInt10 cz1 fz1 pz1 cz2 fz2 pz2;

USEVARIABLES ARE  ID interven cz1 fz1 pz1 cz2 fz2 pz2
BaaT1 BaaE1I BabT2 BabE2I BbaT1 BbaR1I BbbT2 BbbR2I trial;

CLUSTER = ID;

ANALYSIS: TYPE is twolevel basic;
```

Creates new variable for trial. Will match number of columns e.g.:
Baa1-Baa10
Trial = 10
Baa1-Baa20
Trial = 20

Saves data to new file for your MLM with headers indicated in USEVARIABLES.

Output

SUMMARY OF DATA

Number of clusters 33

Size (s)	Cluster ID with Size s
10	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33

.
.
.

SAVEDATA INFORMATION

Save file
Demolong.dat
Order and format of variables

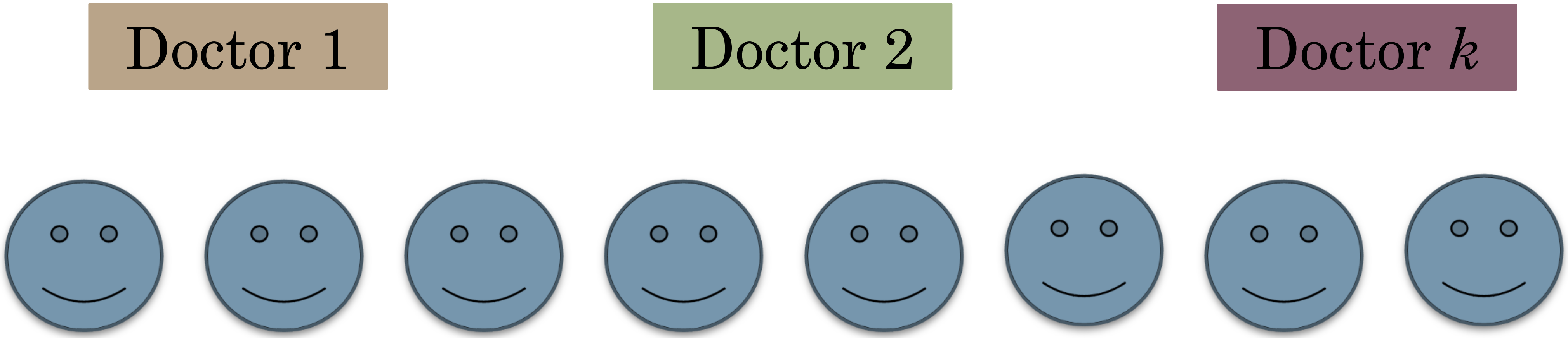
INTERVEN	F10.3
CZ1	F10.3
FZ1	F10.3
PZ1	F10.3
CZ2	F10.3
FZ2	F10.3
PZ2	F10.3
BAAT1	F10.3
BAAE1I	F10.3
BABT2	F10.3
BABE2I	F10.3
BBAT1	F10.3
BBAR1I	F10.3
BBBT2	F10.3
BBBR2I	F10.3
TRIAL	F10.3
ID	I3

Saves data to new file for your MLM with headers indicated in USEVARIABLES.

MLM: Basics

- Regression equation fit at the individual level
- Let parameters of the regression equation vary by group membership
- Use group-level variables to explain variation in the individual-level parameters
- Allows you to test for main effects and interactions at ***within*** and ***between*** levels

EXAMPLE:



Patients

Intraclass Correlation (ICC)

- Model fit – do you have clustering and is it enough clustering?
- ICC is the amount of between-cluster variability relative to the total variation, **intra-cluster homogeneity.**
  - **Rule of thumb:** If ICC is greater than .05 then you should use MLM
  - **In other words:** ICC < .05 = no violation of independence

Intraclass Correlation (ICC)

Resulting in the intraclass correlation:

$$\boldsymbol{\rho}(y_{kj}, y_{lj}) = V(\eta)/[V(\eta) + V(\varepsilon)]$$

Interpretation: It describes how strongly units in the same group resemble each other.

$\boldsymbol{\rho}$: The ICC

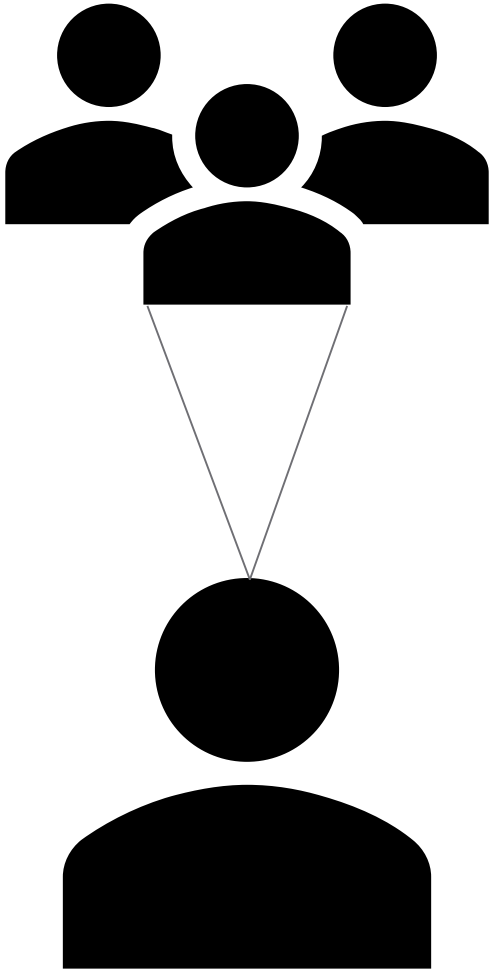
y_{kj}, y_{lj} = k and l are individuals groups

$V(\eta)$ = variance of the random effect

$V(\varepsilon)$ = variability of the error

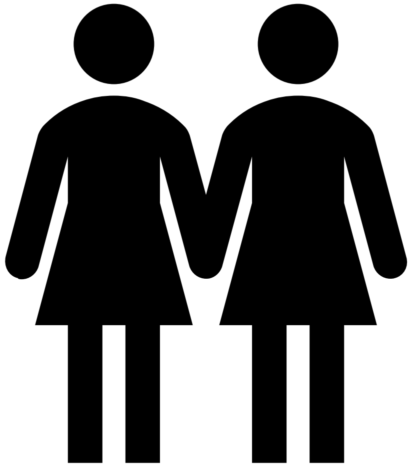
Within = Level 1

- Nested under Level 2
- **Repeated Measures = Individual Level**
- Example variables:
  - **Daily alcohol use** – number of drinks on each day
  - **Weekly depression symptoms** – count of depression symptoms at each session
  - **Job satisfaction ratings each quarter** – summary of a job satisfaction scale collected repeatedly



Between = Level 2

- **Static Variables = Group Level**
  - Sex
  - Ethnicity
  - Treatment condition
  - Baseline substance use (or anything collected one time)



The Level 1 Regression Equation

- Predict value of your DV from the values of your Level 1 IVs
- Equation has the general form of

$$Y_{ij} = B0_j + B1_j\ X1_{ij} + B2_j\ X2_{ij} + r_{ij}$$

The Level 1 Regression Equation

- Predict value of your DV from the values of your Level 1 IVs
- Equation has the general form of

$$Y_{ij} = B0_j + B1_j \; X1_{ij} + B2_j \; X2_{ij} + r_{ij}$$

"i" refers to the person number and "j" refers to the group number

The Level 1 Regression Equation

- Predict value of your DV from the values of your Level 1 IVs
- Equation has the general form of

$$\mathbf{Y_{ij}} = B0_j + B1_j\ \mathbf{X1_{ij}} + B2_j\ \mathbf{X2_{ij}} + \mathbf{r_{ij}}$$

"i" refers to the person number and "j" refers to the group number

Y, X1, X2, and r - vary across individuals **and** groups

The Level 1 Regression Equation

- Predict value of your DV from the values of your Level 1 IVs
- Equation has the general form of

$$Y_{ij} = \mathbf{B0_j} + \mathbf{B1_j}\ X1_{ij} + \mathbf{B2_j}\ X2_{ij} + r_{ij}$$

"i" refers to the person number and "j" refers to the group number

Y, X1, X2, and r - vary across individuals **<u>and</u>** groups

B0, B1, B2 only vary across <u>groups</u>

^this is the variability we can try to explain

Example:

- Y – Math score
- Predictor 1 – IQ
- Predictor 2 – Aptitude test
- Moderator (W) – Teacher's experience
- B0j - the math score when all predictors are zero
- B1j - the association between IQ and math score
- B2j - the association between an aptitude test and math score

Level 2 Equations

- Predict the value of the Level 1 parameters using values of your Level 2 IVs (e.g., W1)
- Sample equations:

$B0_j = \boldsymbol{\gamma}00 + \boldsymbol{\gamma}\ 01\ W1_j + u0_j$

$B1_j = \boldsymbol{\gamma}\ 10 + \boldsymbol{\gamma}\ 11\ \ W1_j + u1_j$

$B2j = \boldsymbol{\gamma}\ 20 + \boldsymbol{\gamma}\ 21\ \ W1_j + u2_j$

- You will have a separate equation for each parameter in the Level 1 equation
  - $\boldsymbol{\gamma}00$ = intercept of the equation predicting the intercept
  - $\boldsymbol{\gamma}10$ = intercept of the equation predicting slope B1
  - $\boldsymbol{\gamma}20$ = intercept of the equation predicting slope B2
  - $\boldsymbol{\gamma}01$ = slope of the equation predicting the intercept
  - $\boldsymbol{\gamma}11$ = slope of the equation predicting slope B1
  - $\boldsymbol{\gamma}21$ = slope of the equation predicting slope B2
  - W1j = predictor variable
  - u0j, u1j, u2j = error terms

Example:
Y – Math score
Predictor 1 – IQ
Predictor 2 – Aptitude test
Moderator (W) – Teacher's experience

Level 2 Equations:

$B0_j = \gamma 00 + \gamma\ 01\ W1_j + u0_j$

$B0_j$ = overall math score (y) when all predictors = 0.

- $\gamma 00$ = intercept of the equation predicting the intercept
  - Overall math score when teacher's experience = 0
- $\gamma 01$ = slope of the equation predicting the intercept
  - The slope of math score moderated by teacher's experience is predicting overall math score
- W1j = moderator variable
- u0j, u1j, u2j = error terms

Level 2 Equations:

Example:
Y – Math score
Predictor 1 – IQ
Predictor 2 – Aptitude test
Moderator (W) – Teacher's experience

$B1_j = \boldsymbol{\gamma}\ 10 + \boldsymbol{\gamma}\ 11\ \ W1_j + u1_j$

$B1_j$ = the slope of the relation between student's IQ and math score, moderated by teacher experience.

- **γ**10 = intercept of the equation predicting slope B1
  - **The intercept predicting the slope of the relation between IQ and math score when experience = 0.**
- **γ**11 = slope of the equation predicting slope B1
  - **Slope predicting the relation between IQ and math score.**
  - **Positive indicates steeper slope**
- W1j = moderator variable
- u0j, u1j, u2j = error terms

Level 2 Equations:

Example:
Y – Math score
Predictor 1 – IQ
Predictor 2 – Aptitude test
Moderator (W) – Teacher's experience

$B2j = \boldsymbol{\gamma}\ 20 + \boldsymbol{\gamma}\ 21\ \ W1_j + u2_j$

B2j = the relation between aptitude test and math score, moderated by teacher's experience.

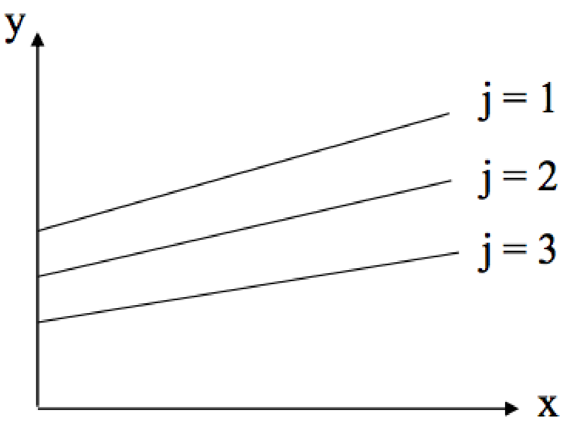
- $\boldsymbol{\gamma}20$ = intercept of the equation predicting slope B2
  - **The intercept predicting the slope of the relation between aptitude and math score when experience = 0.**
- $\boldsymbol{\gamma}21$ = slope of the equation predicting slope B2
  - **Slope predicting the relation between aptitude and math score.**
  - **Positive indicates steeper slope**
- W1j = moderator variable
- u0j, u1j, u2j = error terms

Visualizing MLM

Within level

Individual i in cluster j

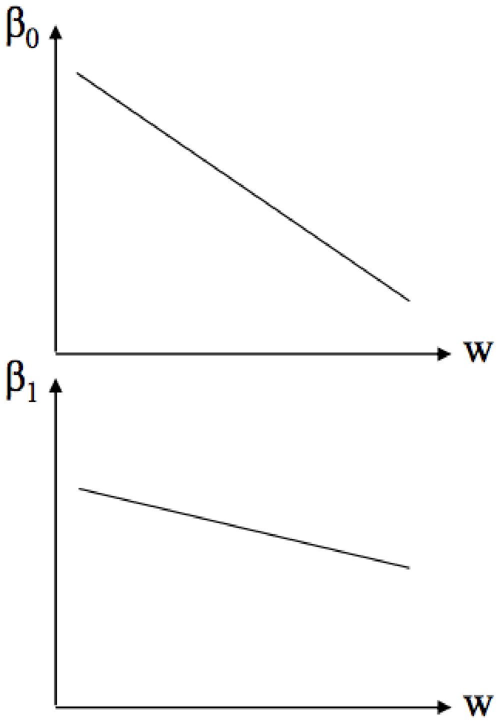
(1) $y_{ij} = \beta_{0j} + \beta_{1j}\, x_{ij} + r_{ij}$



Between level

(2a) $\beta_{0j} = \gamma_{00} + \gamma_{01}\, w_j + u_{0j}$

(2b) $\beta_{1j} = \gamma_{10} + \gamma_{11}\, w_j + u_{1j}$



Two-level analysis (individual i in cluster j):

y_{ij} : individual-level outcome variable

x_{ij} : individual-level covariate

w_j : cluster-level covariate

Random intercepts, random slopes:

Level 1 (*Within*) : $y_{ij} = \beta_{0j} + \beta_{1j}\, x_{ij} + r_{ij}$,

Level 2 (*Between*) : $\beta_{0j} = \gamma_{00} + \gamma_{01}\, w_j + u_{0j}$,

Level 2 (*Between*) : $\beta_{1j} = \gamma_{10} + \gamma_{11}\, w_j + u_{1j}$.

The Combined Model

$$Y_{ij} = B0_j + B1_j\ X1_{ij} + B2_j\ X2_{ij} + r_{ij}$$

$$\gamma 00 + \gamma\ 01\ W1_j + u0_j \qquad \gamma\ 10 + \gamma\ 11\ W1_j + u1_j \qquad \gamma\ 20 + \gamma\ 21\ W1_j + u2_j$$

We can substitute the Level 2 equations into the Level 1 equation to see the combined model

The Combined Model

$$Y_{ij} = \gamma 00 + \gamma\ 01\ W1_j + u0_j + \gamma\ 10 + \gamma\ 11\ \ W1_j + u1_j X1_{ij} + \gamma\ 20 + \gamma\ 21\ \ W1_j + u2_j X2_{ij} + r_{ij}$$

We can substitute the Level 2 equations into the Level 1 equation to see the combined model

Cannot estimate this using normal regression

Centering

- Centering is not necessary, depends on interpretation goals
  - see references in dropbox
- **Level 1 regression equation:**

$$Y_{ij} = B0_j + B1_j * X1_{ij} + B2_j * X2_{ij} + r_{ij}$$

- $B0_j$ tells us the value of Y_{ij} when $X1_{ij} = 0$ and $X2_{ij} = 0$
  - i.e.: $Y_{ij} = B0_j \cancel{+ B1_j * X1_{ij} + B2_j * X2_{ij}} + r_{ij}$
- Interpretation of $B0_j$ depends on the **scale** of $X1_{ij}$ and $X2_{ij}$
- “Centering” refers to subtracting a value from an X to make the 0 point meaningful:
  - Thus, if both X1 and X2 are group mean centered then $B0_j$ tells us the value of Y_{ij} when $X1_{ij}$ and $X2_{ij}$ are at their group means

Group Mean Centering

- Within cluster/person
  - Only possible on Level 1
  - Calculate deviation of each person's score with the mean of the cluster to which the person belongs

$$Stress_{centered} = Stress_{it} - \overline{X}_{stress_t}$$

  - Longitudinal case - calculate deviation of each person's score with his/her mean level

Grand Mean Centering

- Between cluster/person
  - Calculate deviation of each score from the mean score (across clusters/persons)

$$Stress_{centered} = Stress_{it} - \overline{X}_{stress}$$

  - Thus, averaging across clusters. The Level 1 predictor includes within and between cluster variance

Centering recommendations from Enders & Tofighi (2007)

- Level 2 center at grand mean

- Level 1 centering depends on research question…

    1. Association between Level 1 IVs and DVs is primary interest – group mean center
    2. Level 2 predictor is primary interest and want to control for Level 1 covariates – grand mean center
    3. Cross-level interactions or interactions on Level 1 are primary interest – group mean center

Fixed vs. Random Effects

- **Fixed:**
  - Has only a single value in the model and is applied to ***each level-1 unit*** in the analysis regardless of the level-2 unit under which a case is nested.
  - Measured without error
  - **Examples:** gender, treatment group
- **Random:**
  - **Varies** between Level-**2** units.
  - ***Think about this as free to change (as opposed to due to chance).***
  - Measured with error, values come from a larger population of possible values
  - **Examples:** mood across time, scores within a classroom

The Significance of Fixed and Random Effects

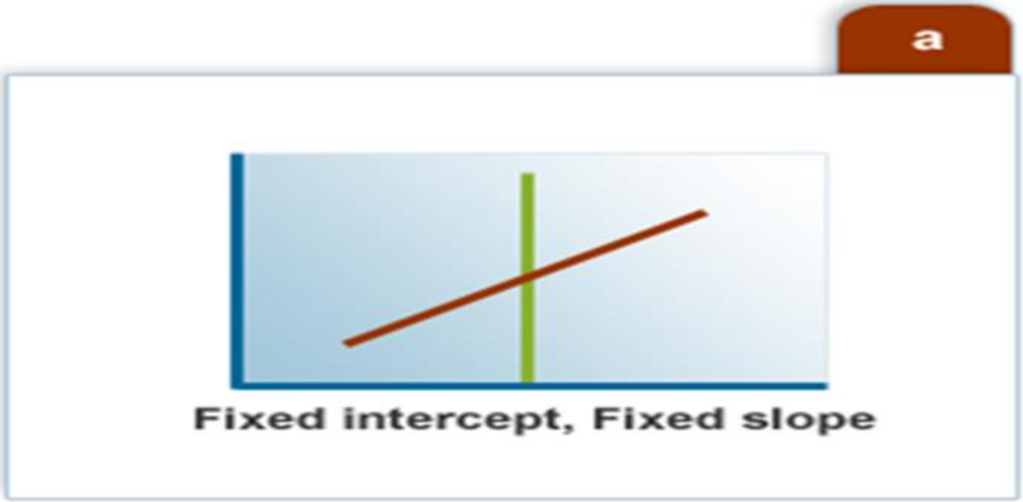
- **<u>Ordinary regression:</u>**
  - Fixed: Regression intercept and regression coefficient
  - Random: Regression residual
    - Individual deviation from their predicted value
- **<u>MLM:</u>**
  - By estimating 1 or all coefficients as random, you can test:
    - Several regression intercepts for the model
    - Several regression coefficients for each predictor
      - One for each level-2 unit

How to Choose Random or Fixed

- Intercepts are commonly estimated as random effects
  - Mean centering: *Means are different between Level-2 units*.
- Theoretically based:
  - If the relationship between level-1 variables and the outcome differs between level-2 units, this suggests setting the effect as random.

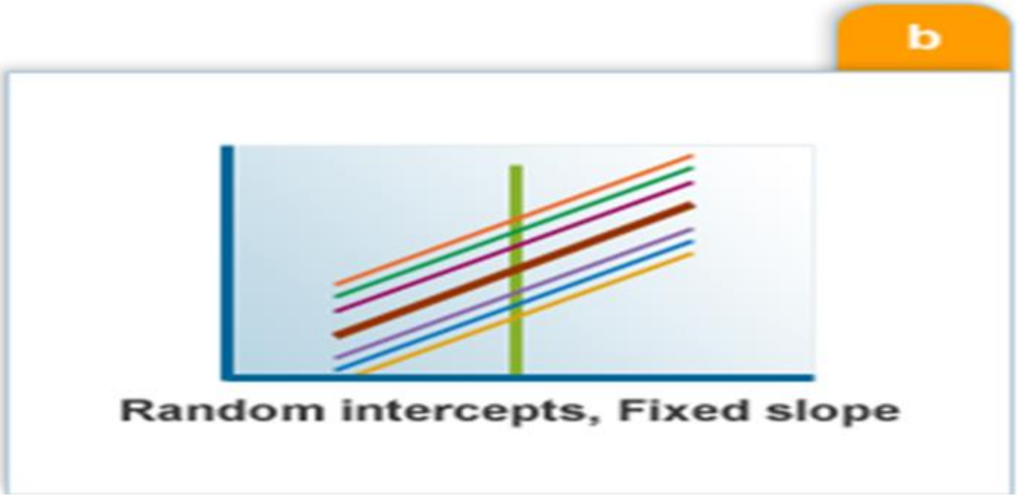
<u>Fixed intercept, Fixed slope:</u>

- Both the intercept and the slope are aggregated to a straight line.
- No ability to vary



<u>Random intercept, Fixed slope:</u>

- Allow intercept to vary
  - Intercept can take many patterns
- Slope is the same across groups



Random intercept, Random slope

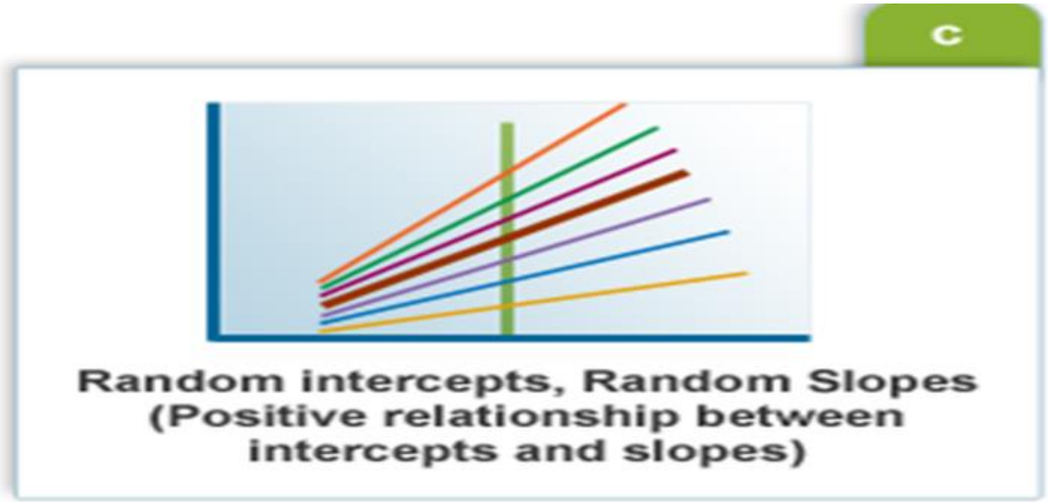
Both intercept and slope are free to vary across groups.

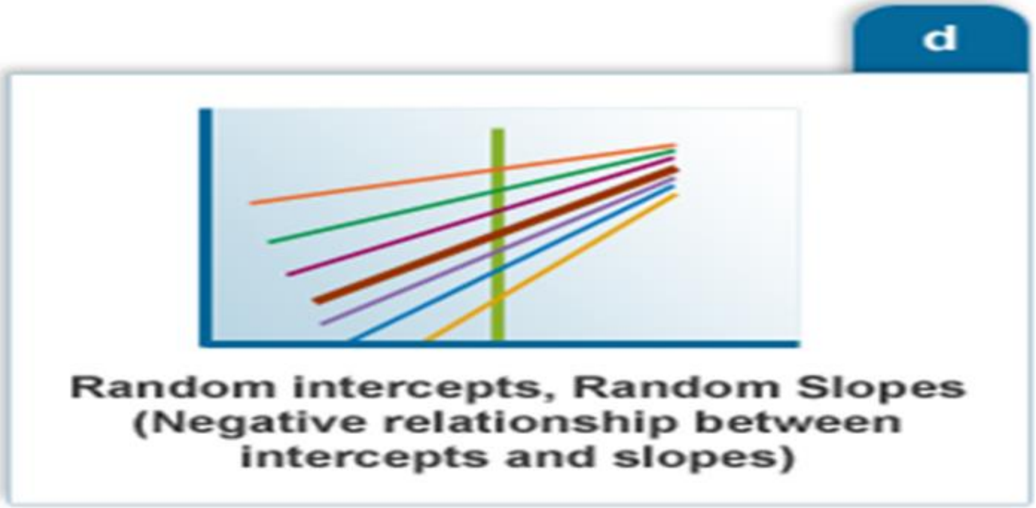
<u>Positive Relationship:</u>

- Steeper slopes = stronger relationship
- Flatter slopes = weaker relationship

<u>Negative Relationship:</u>

- Steeper slopes = stronger relationship
- Flatter slopes = weaker relationship





No Relationship

Both intercepts and slopes are free to vary, but no relationship is evident

Random intercepts, Random Slopes (No relationship between intercepts and slopes)

Example:

Alcohol consequences among college students by PTSD status

- N = 600 college students
- Screened for PTSD
  - Three groups: Criterion A- (i.e., no trauma), PTSD+, PTSD-
- Daily reporting of alcohol related consequences for 30 days
  - 17985 observations
- Pregaming (PGDAYS) also assessed on each day (0 = did not pregame; 1 = did pregame)

Analysis Plan: General

A series of multilevel models were conducted to test the following study hypotheses:

H1: College student's alcohol related consequences (ARC) vary across 30 days.
Filename: PG PTSD12 int only.out
H2: Student's ARCs are higher on days when they drank alcohol before going out (e.g., pregaming)
Filename: PG PTSD rand coef 1 predictor.out
H3: Having a history of trauma moderated the associated between pregaming and ARC.
Filename: PG PTSD rand coef trauma vs no trauma.out

ARC was treated as a continuous normally distributed variable. Pregaming was coded 0 = did not pregame, 1 = did pregame. Trauma was coded 0 = Trauma, 1 = No trauma. Analyses were conducted using Mplus 7.4 (Muthén & Muthén, 1998–2012). ARC was treated as both a within and between level variable as it was expected to vary from day to day as well as across individuals. Pregaming was treated as a within level only variable, as each day was coded as either a day in which a student pregamed or did not pregame. Trauma was coded as a between level only variable as trauma exposure was established at baseline.

Hypothesis 1: Random Intercept Only

H1: College students' alcohol related consequences (ARC) vary across 30 days

- Filename: PG PTSD12 int only.out

Analysis Plan (from write up):

Hypothesis 1 was tested with an intercept only model by examining the intraclass correlation (ICC). The ICC is the proportion of variance in the outcome variable that is explained by the grouping structure of the hierarchical model. Values higher than .05 are typically considered sufficient to necessitate multilevel modeling due to alpha inflation that results from dependency in nested data.

Random Intercept Model:

FILENAME: PGPTSD12 INT ONLY INT.INP

Number of alcohol related consequences reported per day ("CONSd")

$$\boxed{\begin{aligned} CONSd_{ti} &= \beta_{oi} + e_{ti} \\ \beta_{oi} &= \gamma_{oo} + u_{oi} \end{aligned}}$$

```
USEVARIABLES ARE
CONSd      ;

CLUSTER IS IDNUM;

ANALYSIS:
TYPE IS TWOLEVEL RANDOM;

MODEL:
%WITHIN%
consd;

%BETWEEN%
consd;
```

Random Intercept Model:

Number of alcohol related consequences in the past 30 days ("CONSd")

$$CONSd_{ti} = \beta_{oi} + e_{ti}$$
$$\beta_{oi} = \gamma_{oo} + u_{oi}$$

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Within Level

 Variances
    CONSD              1.292      0.099     12.997      0.000

Between Level

 Means
    CONSD              0.319      0.016     19.537      0.000

 Variances
    CONSD              0.116      0.031      3.776      0.000
```

Level 1 residual variance
Deviation of each day of CONSd
From the student's mean number of CONSd

Mean CONSd for the ith student across 30 days

Level 2 Variance
Individuals differences from the grand mean

Intraclass Correlation Coefficient Example

- Index of the degree of dependency in the data
  - Proportion of total variance account for by Level 2
    - 8% of variance in alcohol consequences is due to between-person
    - 92% within-individuals

$$ICC = \frac{Level2\text{variance}}{Level2\text{variance} + Level1\text{variance}}$$

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Within Level

 Variances
    CONSD              1.292      0.099     12.997      0.000

Between Level

 Means
    CONSD              0.319      0.016     19.537      0.000

 Variances
    CONSD              0.116      0.031      3.776      0.000
```

$$ICC = \frac{.116}{(.116 + 1.292)} = .082$$

```
Estimated Intraclass Correlations for the Y Variables

             Intraclass
Variable     Correlation

CONSD           0.082
```

Results:

Individual level variability in ARC. The ICC for the intercept only model was .08, which indicated that 8% of the variance was explained by dependencies across individuals. This indicates sufficient individual level variability in ARC to warrant MLM. Further, the mean ARC across students was .32, which suggests that an average student reports a third of a consequence on a given day. In other words, on most days, students did not report any consequences .

Results: (Paragraph 1 in Write Up)

Individual level variability in ARC. The ICC for the intercept only model was .08, which indicated that 8% of the variance was explained by dependencies across individuals. This indicates sufficient individual level variability in ARC to warrant MLM. Further, the mean ARC across students was .32, which suggests that an average student reports a third of a consequence on a given day. In other words, on most days, students did not report any consequences .

Hypothesis 2: Level-1 Predictor Model

Analysis Plan: (From write-up)

H2: Student's ARCs are higher on days when they drank alcohol before going out (e.g., pregaming)

Filename: PG PTSD rand coef 1 predictor.out

[…]

Hypothesis 2 was tested by adding pregaming to the model on the within level as a predictor of ARC.

Random Coefficient Model with Level 1 Predictor

FILENAME: PG PTSD rand coef 1 predictor.INP

VARIABLE:

USEVARIABLES ARE CONSd PGDAY;

CLUSTER IS IDNUM;

WITHIN IS PGDAY;

ANALYSIS:

TYPE IS TWOLEVEL RANDOM;

MODEL:

%WITHIN%

s | CONSd on PGDAY;

%BETWEEN%

CONSd;

s;

CONSd with s;

$$\boxed{\begin{aligned} &CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti} \\ &\beta_{oi} = \gamma_{oo} + u_{oi} \\ &\beta_{1i} = \gamma_{1o} + u_{1i} \end{aligned}}$$

This creates a random slope

Random Coefficient Model with Level 1 Predictor

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo} + u_{oi}$$
$$\beta_{1i} = \gamma_{1o} + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.929	0.000
Between Level				
CONSD WITH				
S	0.153	0.039	3.870	0.000
Means				
CONSD	0.184	0.011	16.563	0.000
S	2.239	0.114	19.674	0.000
Variances				
CONSD	0.044	0.012	3.525	0.000
S	4.463	0.777	5.747	0.000

Intercept – consd = .184 = mean number of consequences for the ith individual across the 30 days when pgday = 0

Random Coefficient Model with Level 1 Predictor

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo} + u_{oi}$$
$$\beta_{1i} = \gamma_{1o} + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.929	0.000
Between Level				
CONSD WITH				
S	0.153	0.039	3.870	0.000
Means				
CONSD	0.184	0.011	16.563	0.000
S	2.239	0.114	19.674	0.000
Variances				
CONSD	0.044	0.012	3.525	0.000
S	4.463	0.777	5.747	0.000

<u>Intercept</u> – consd = .184 = mean number of consequences for the ith individual across the 30 days when pgday = 0

<u>Within slope (consd on pgday)</u> = 2.239 – linear relationship between pgday and consd within individuals

Random Coefficient Model with Level 1 Predictor

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo} + u_{oi}$$
$$\beta_{1i} = \gamma_{1o} + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.929	0.000
Between Level				
CONSD WITH				
S	0.153	0.039	3.870	0.000
Means				
CONSD	0.184	0.011	16.563	0.000
S	2.239	0.114	19.674	0.000
Variances				
CONSD	0.044	0.012	3.525	0.000
S	4.463	0.777	5.747	0.000

Intercept – consd = .184 = mean number of consequences for the ith individual across the 30 days when pgday = 0

Within slope (consd on pgday) = 2.239 – linear relationship between pgday and consd within individuals

Level 1 Residual = .832 = deviation of each student's consd from the student's mean consd

Random Coefficient Model with Level 1 Predictor

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo} + u_{oi}$$
$$\beta_{1i} = \gamma_{1o} + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.929	0.000
Between Level				
CONSD WITH				
S	0.153	0.039	3.870	0.000
Means				
CONSD	0.184	0.011	16.563	0.000
S	2.239	0.114	19.674	0.000
Variances				
CONSD	0.044	0.012	3.525	0.000
S	4.463	0.777	5.747	0.000

Intercept – consd = .184 = mean number of consequences for the ith individual across the 30 days when pgday = 0

Within slope (consd on pgday) = 2.239 – linear relationship between pgday and consd within individuals

Level 1 Residual = .832 = deviation of each student's consd from the student's mean consd

Level 2 intercept variance = .044 = individual differences in consd at pgday = 0

Random Coefficient Model with Level 1 Predictor

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo} + u_{oi}$$
$$\beta_{1i} = \gamma_{1o} + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.929	0.000
Between Level				
CONSD WITH				
S	0.153	0.039	3.870	0.000
Means				
CONSD	0.184	0.011	16.563	0.000
S	2.239	0.114	19.674	0.000
Variances				
CONSD	0.044	0.012	3.525	0.000
S	4.463	0.777	5.747	0.000

Intercept – consd = .184 = mean number of consequences for the ith individual across the 30 days when pgday = 0

Within slope (consd on pgday) = 2.239 – linear relationship between pgday and consd within individuals

Level 1 Residual = .832 = deviation of each student's consd from the student's mean consd

Level 2 intercept variance = .044 = individual differences in consd at pgday = 0

Level 2 slope variance = 4.463 = indiv differences in linear relationship between pgday and consd

Random Coefficient Model with Level 1 Predictor

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo} + u_{oi}$$
$$\beta_{1i} = \gamma_{1o} + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.929	0.000
Between Level				
CONSD WITH				
S	0.153	0.039	3.870	0.000
Means				
CONSD	0.184	0.011	16.563	0.000
S	2.239	0.114	19.674	0.000
Variances				
CONSD	0.044	0.012	3.525	0.000
S	4.463	0.777	5.747	0.000

Intercept – consd = .184 = mean number of consequences for the ith individual across the 30 days when pgday = 0

Within slope (consd on pgday) = 2.239 – linear relationship between pgday and consd within individuals

Level 1 Residual = .832 = deviation of each student's consd from the student's mean consd

Level 2 intercept variance = .044 = individual differences in consd at pgday = 0

Level 2 slope variance = 4.463 = indiv differences in linear relationship between pgday and consd

Level 2 covariance = .153 = association between intercept and slope

Random Coefficient Model with Level 1 Predictor

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo} + u_{oi}$$
$$\beta_{1i} = \gamma_{1o} + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.929	0.000
Between Level				
CONSD WITH				
S	0.153	0.039	3.870	0.000
Means				
CONSD	0.184	0.011	16.563	0.000
S	2.239	0.114	19.674	0.000
Variances				
CONSD	0.044	0.012	3.525	0.000
S	4.463	0.777	5.747	0.000

Level 2 covariance = .153 = association between intercept and slope

- Significant and positive, indicating that individuals who had more consequences when pgday = 0 tended to have increases in consequences on pgdays

Results: (Paragraph 2 in Write up)

Pregaming as a predictor of ARC. Results from the random intercepts and slopes model with pregaming as a level-1 predictor revealed that pregaming is a significant predictor of ARC across individuals (s = 2.39, SE = .11, $p < .001$). Specifically, students reported 2.39 more consequences on pregaming days compared to days when they did not pregame.

Hypothesis 3: Level-1 & Level- 2 Predictor Model

Analysis Plan: (From write-up)

H3: Having a history of trauma moderated the associated between pregaming and ARC.

Filename: PG PTSD rand coef trauma vs no trauma.out

[…]

Hypothesis 3 was tested by adding trauma to the model as a predictor of the random intercept and random slope terms.

Random Coefficient Model with Level 1 and Level 2 Predictors

FILENAME: PG PTSD rand coef trauma vs no trauma.INP

```
VARIABLE:
  USEVARIABLES ARE CONSd PGDAY PTSD1;
  CLUSTER IS IDNUM;
  WITHIN IS pgday;
  BETWEEN IS ptsd1;
ANALYSIS:
   TYPE IS TWOLEVEL RANDOM;
MODEL:
   %WITHIN%
   s | consd on pgday;
   %BETWEEN%
   consd s on ptsd1;
   consd with s;
```

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo}\,(PTSD1) + u_{oi}$$
$$\beta_{1i} = \gamma_{1o}\,(PTSD1) + u_{1i}$$

Random Coefficient Model with Level 1 and 2 Predictors

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo}\,(PTSD1) + u_{oi}$$
$$\beta_{1i} = \gamma_{1o}\,(PTSD1) + u_{1i}$$

Intercept – consd = .226 = mean number of consequences for the ith individual across the 30 days when pgday = 0 at the unweighted mean of ptsd groups

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Within Level

 Residual Variances
    CONSD              0.832      0.070     11.930      0.000

Between Level

 S          ON
    PTSD1             -0.522      0.230     -2.271      0.023

 CONSD      ON
    PTSD1             -0.106      0.020     -5.411      0.000

 CONSD    WITH
    S                  0.139      0.037      3.780      0.000

 Intercepts
    CONSD              0.226      0.017     13.558      0.000
    S                  2.436      0.145     16.783      0.000

 Residual Variances
    CONSD              0.041      0.012      3.478      0.001
    S                  4.400      0.772      5.699      0.000
```

Random Coefficient Model with Level 1 and 2 Predictors

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo}\,(PTSD1) + u_{oi}$$
$$\beta_{1i} = \gamma_{1o}\,(PTSD1) + u_{1i}$$

Intercept – consd = .226 = mean number of consequences for the ith individual across the 30 days when pgday = 0 at the unweighted mean of ptsd groups

Intercept - S = 2.436 – linear relationship between pgday and consd within individuals

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Within Level

 Residual Variances
    CONSD              0.832      0.070     11.930      0.000

Between Level

 S          ON
    PTSD1             -0.522      0.230     -2.271      0.023

 CONSD      ON
    PTSD1             -0.106      0.020     -5.411      0.000

 CONSD    WITH
    S                  0.139      0.037      3.780      0.000

 Intercepts
    CONSD              0.226      0.017     13.558      0.000
    S                  2.436      0.145     16.783      0.000

 Residual Variances
    CONSD              0.041      0.012      3.478      0.001
    S                  4.400      0.772      5.699      0.000
```

Random Coefficient Model with Level 1 and 2 Predictors

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo}\,(PTSD1) + u_{oi}$$
$$\beta_{1i} = \gamma_{1o}\,(PTSD1) + u_{1i}$$

Intercept – consd = .226 = mean number of consequences for the ith individual across the 30 days when pgday = 0 at the unweighted mean of ptsd groups

Intercept - S = 2.436 – linear relationship between pgday and consd within individuals

Consd on PTSD = -.106 – effect of trauma on consd at pgday=0 (negative means no trauma = less consequences)

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.930	0.000
Between Level				
S ON				
PTSD1	-0.522	0.230	-2.271	0.023
CONSD ON				
PTSD1	-0.106	0.020	-5.411	0.000
CONSD WITH				
S	0.139	0.037	3.780	0.000
Intercepts				
CONSD	0.226	0.017	13.558	0.000
S	2.436	0.145	16.783	0.000
Residual Variances				
CONSD	0.041	0.012	3.478	0.001
S	4.400	0.772	5.699	0.000

Random Coefficient Model with Level 1 and 2 Predictors

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo}\,(PTSD1) + u_{oi}$$
$$\beta_{1i} = \gamma_{1o}\,(PTSD1) + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.930	0.000
Between Level				
S ON				
PTSD1	-0.522	0.230	-2.271	0.023
CONSD ON				
PTSD1	-0.106	0.020	-5.411	0.000
CONSD WITH				
S	0.139	0.037	3.780	0.000
Intercepts				
CONSD	0.226	0.017	13.558	0.000
S	2.436	0.145	16.783	0.000
Residual Variances				
CONSD	0.041	0.012	3.478	0.001
S	4.400	0.772	5.699	0.000

<u>Intercept – consd</u> = .226 = mean number of consequences for the ith individual across the 30 days when pgday = 0 at the unweighted mean of ptsd groups

<u>Intercept - S</u> = 2.436 – linear relationship between pgday and consd within individuals

<u>Consd on PTSD</u> = -.106 – effect of trauma on consd at pgday=0 (negative means no trauma = less consequences)

<u>S on PTSD</u> = -.522 – cross level interaction – individuals with a trauma history had steeper slope between pgday and consd

Random Coefficient Model with Level 1 and 2 Predictors

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo}\,(PTSD1) + u_{oi}$$
$$\beta_{1i} = \gamma_{1o}\,(PTSD1) + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.930	0.000
Between Level				
S ON				
PTSD1	-0.522	0.230	-2.271	0.023
CONSD ON				
PTSD1	-0.106	0.020	-5.411	0.000
CONSD WITH				
S	0.139	0.037	3.780	0.000
Intercepts				
CONSD	0.226	0.017	13.558	0.000
S	2.436	0.145	16.783	0.000
Residual Variances				
CONSD	0.041	0.012	3.478	0.001
S	4.400	0.772	5.699	0.000

<u>Intercept – consd</u> = .226 = mean number of consequences for the ith individual across the 30 days when pgday = 0 at the unweighted mean of ptsd groups

<u>Intercept - S</u> = 2.436 – linear relationship between pgday and consd within individuals

<u>Consd on PTSD</u> = -.106 – effect of trauma on consd at pgday=0 (negative means no trauma = less consequences)

<u>S on PTSD</u> = -.522 – cross level interaction – individuals with a trauma history had steeper slope between pgday and consd

<u>Level 1 Residual</u> = .832 = deviation of each day of consequences from the individual's mean consd

Random Coefficient Model with Level 1 and 2 Predictors

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo}\,(PTSD1) + u_{oi}$$
$$\beta_{1i} = \gamma_{1o}\,(PTSD1) + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.930	0.000
Between Level				
S ON				
PTSD1	-0.522	0.230	-2.271	0.023
CONSD ON				
PTSD1	-0.106	0.020	-5.411	0.000
CONSD WITH				
S	0.139	0.037	3.780	0.000
Intercepts				
CONSD	0.226	0.017	13.558	0.000
S	2.436	0.145	16.783	0.000
Residual Variances				
CONSD	0.041	0.012	3.478	0.001
S	4.400	0.772	5.699	0.000

Intercept – consd = .226 = mean number of consequences for the ith individual across the 30 days when pgday = 0 at the unweighted mean of ptsd groups

Intercept - S = 2.436 – linear relationship between pgday and consd within individuals

Consd on PTSD = -.106 – effect of trauma on consd at pgday=0 (negative means no trauma = less consequences)

S on PTSD = -.522 – cross level interaction – individuals with a trauma history had steeper slope between pgday and consd

Level 1 Residual = .832 = deviation of each day of consequences from the individual's mean consd

Level 2 intercept variance = .041 = individual differences in consd at pgday = 0

Random Coefficient Model with Level 1 and 2 Predictors

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo}\,(PTSD1) + u_{oi}$$
$$\beta_{1i} = \gamma_{1o}\,(PTSD1) + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.930	0.000
Between Level				
S ON				
PTSD1	-0.522	0.230	-2.271	0.023
CONSD ON				
PTSD1	-0.106	0.020	-5.411	0.000
CONSD WITH				
S	0.139	0.037	3.780	0.000
Intercepts				
CONSD	0.226	0.017	13.558	0.000
S	2.436	0.145	16.783	0.000
Residual Variances				
CONSD	0.041	0.012	3.478	0.001
S	4.400	0.772	5.699	0.000

Intercept – consd = .226 = mean number of consequences for the ith individual across the 30 days when pgday = 0 at the unweighted mean of ptsd groups

Intercept - S = 2.436 – linear relationship between pgday and consd within individuals

Consd on PTSD = -.106 – effect of trauma on consd at pgday=0 (negative means no trauma = less consequences)

S on PTSD = -.522 – cross level interaction – individuals with a trauma history had steeper slope between pgday and consd

Level 1 Residual = .832 = deviation of each day of consequences from the individual's mean consd

Level 2 intercept variance = .041 = individual differences in consd at pgday = 0

Level 2 slope variance = 4.4 = individual differences in linear relationship between pgday and consd at the unweighted mean of ptsd1

Random Coefficient Model with Level 1 and 2 Predictors

$$CONSd_{ti} = \beta_{oi} + \beta_{oi}(PGDAY) + e_{ti}$$
$$\beta_{oi} = \gamma_{oo}\,(PTSD1) + u_{oi}$$
$$\beta_{1i} = \gamma_{1o}\,(PTSD1) + u_{1i}$$

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
CONSD	0.832	0.070	11.930	0.000
Between Level				
S ON				
PTSD1	-0.522	0.230	-2.271	0.023
CONSD ON				
PTSD1	-0.106	0.020	-5.411	0.000
CONSD WITH				
S	0.139	0.037	3.780	0.000
Intercepts				
CONSD	0.226	0.017	13.558	0.000
S	2.436	0.145	16.783	0.000
Residual Variances				
CONSD	0.041	0.012	3.478	0.001
S	4.400	0.772	5.699	0.000

Intercept – consd = .226 = mean number of consequences for the ith individual across the 30 days when pgday = 0 at the unweighted mean of ptsd groups

Intercept - S = 2.436 – linear relationship between pgday and consd within individuals

Consd on PTSD = -.106 – effect of trauma on consd at pgday=0 (negative means no trauma = less consequences)

S on PTSD = -.522 – cross level interaction – individuals with a trauma history had steeper slope between pgday and consd

Level 1 Residual = .832 = deviation of each day of consequences from the individual's mean consd

Level 2 intercept variance = .041 = individual differences in consd at pgday = 0

Level 2 slope variance = 4.4 = individual differences in linear relationship between pgday and consd at the unweighted mean of ptsd1

Level 2 covariance = .139 = association between intercept and slope – significant and positive, indicating that individuals who had more consd when pgday = 0 at the unweighted mean of ptsd1 tended to have increases in consequences on pgdays

Results: (Paragraph 3 in Write Up)

Past trauma as a moderator of the pregaming-ARC relationship. Results from the random intercepts and slopes model with pregaming as a level-1 predictor of ARC and trauma history as a level-2 predictor of the slope of pregaming-ARC, revealed that trauma history significantly moderated the pregaming-ARC relationship ($b = -.522$, $SE = .23$, $p = .02$). Specifically, those with a history of trauma had a stronger relationship (i.e., a steeper slope) between pregaming and ARC. Further, having a trauma history was associated with more consequences on level-2, such that having a history of trauma was associated with experiencing more consequences on average compared to students who did not experience trauma ($b = -.11$, $SE = .02$, $p < .001$).

Discussion: (Write Up)

Random Intercept Model

The present study demonstrated that there is significant variation in ARC within persons in a sample of college students. In addition, this present study demonstrated that students report more consequences on days when they pregame compared to days when they do not pregame. Finally, this relationship was moderated by trauma history, such that those with a history of trauma experienced more consequences on average and had a stronger relationship between pregaming and ARC. Intervention efforts aimed at reducing ARC among college students should target limiting pregaming, and special attention should be paid to students who report a history of trauma, as trauma portends additional risk for experiencing ARC.

Discussion: (Write Up)

The present study demonstrated that there is significant variation in ARC within persons in a sample of college students. In addition, this present study demonstrated that students report more consequences on days when they pregame compared to days when they do not pregame. Finally, this relationship was moderated by trauma history, such that those with a history of trauma experienced more consequences on average and had a stronger relationship between pregaming and ARC. Intervention efforts aimed at reducing ARC among college students should target limiting pregaming, and special attention should be paid to students who report a history of trauma, as trauma portends additional risk for experiencing ARC.

Random Coefficient Model with Level 1 Predictor

Discussion: (Write Up)

The present study demonstrated that there is significant variation in ARC within persons in a sample of college students. In addition, this present study demonstrated that students report more consequences on days when they pregame compared to days when they do not pregame. Finally, this relationship was moderated by trauma history, such that those with a history of trauma experienced more consequences on average and had a stronger relationship between pregaming and ARC. Intervention efforts aimed at reducing ARC among college students should target limiting pregaming, and special attention should be paid to students who report a history of trauma, as trauma portends additional risk for experiencing ARC.

Random Coefficient Model with Level 1& Level 2 Predictors

Summary:
How to write up the results

- In the analysis plan
  - Describe the data structure
    - Normality, clustering, centering, etc.
  - Model building procedure and decisions
    - The three models we discussed are nested so we can use BICs for comparisons
    - Two models are nested if both contain the same terms and one has at least one additional term. + ϵ (2) Model (1) is nested within model (2). Model (1) is the reduced model and model (2) is the full model.
      - Intercept only nested within random slope nested within 1 predictor nested with 2 predictor
- In the results section
  - Only report on the final model
  - Overall fit
  - ICCs
  - Key parameters
- Discussion
  - Focus on within and between inferences