Three-Level Longitudinal Models for Ecological Momentary Assessment Data

- Topics:
  - **Example 6: Predicting fatigue in EMA data**
  - More general info about three-level models

What determines the number of levels?

- **Answer: the model for the “outcome” variance (but keep in mind that any variable can be an “outcome” in multivariate MLM / SEM / M-SEM)**
- How many dimensions of sampling in each outcome?
  - Once per day per person? → 2-level model
  - Multiple times per day per person? → 3-level model
  - What’s the difference? Whether there is an extra correlation of the residuals for observations collected in the same day
- What about the predictors? Their variance still matters!
  - Need to know at which levels they contain variability
  - Is a logical precursor as to the levels at which they can have relationships with other variables (at that level)

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

Notation: t = level-1 time, d = level-2 day, i = level-3 person

Level 1: $y_{tdi} = \beta_{0di} + e_{tdi}$

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

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

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

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

Fixed Intercept =grand mean (because no predictors yet)

Person Random Intercept = person-specific deviation from fixed intercept

3 Total Parameters:

Model for the Means (1):

- Fixed Intercept γ_{00}

Model for the Variance (2):

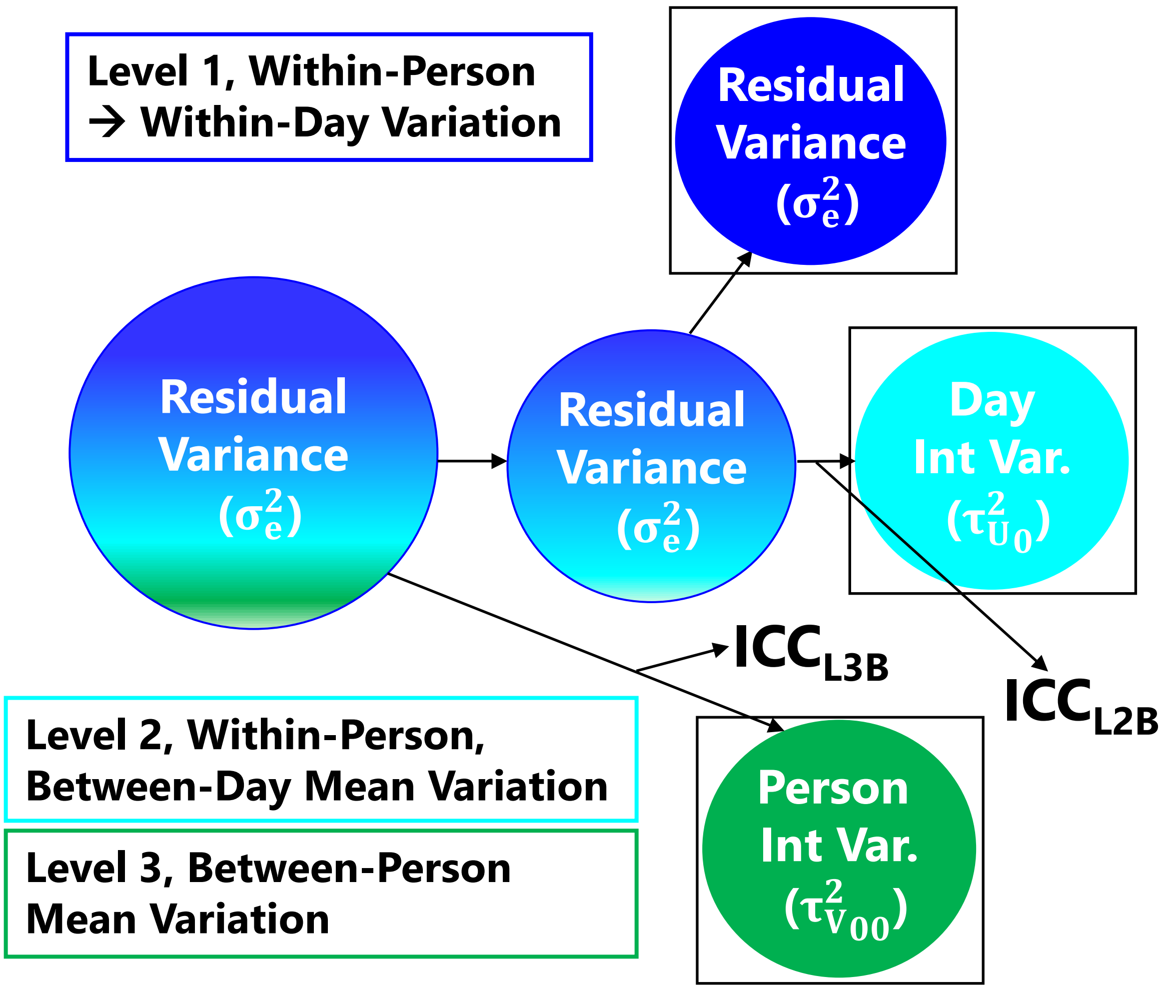
- Level-1 Variance of e_{tdi} → σ_e^2
- Level-2 Variance of U_{0di} → $\tau_{U_0}^2$
- Level-3 Variance of V_{00i} → $\tau_{V_{00}}^2$

Composite equation:

$y_{tdi} = \gamma_{000} + V_{00i} + U_{0d} + e_{tdi}$

Btw: My bad for reusing "V"

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



Useful ICC variants for this type of design:

ICC_{L3B} = L3 / total

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

ICC_{L2B} = L2 / [L2+L1]

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

Example 6: Affect, Health, and Behavior Study (PI: Chris Cushing, KU)

- 25 adolescents completed surveys about their mood and energy levels up to four times per day for up to 20 days (total *N* = 976); only days 3–20 used here to minimize measurement reactivity
  - Fatigue: measured four times per day using 3 items (each response from 1–5) → time-level outcome
  - Negative Affect: measured four times per day using 5 items (each response from 1–5) → time-level predictor
  - Hours of Sleep: previous night's sleep → day-level moderator
- Predictors related to time:
  - Time of day: exact time of within-day observation
  - Day of study and day of week
- For a complete example using SAS, see Example 8b from this class
- First step: ICCs and time-related trends per variable…

Within-Day Fatigue: 2 levels or 3?

Level 1 Time: $\text{Fatigue}_{tdi} = \beta_{0di} + e_{tdi}$

Level 2 Day: Intercept: $\beta_{0di} = \delta_{00i} + 0$

Level 3 Person: Intercept: $\delta_{00i} = \gamma_{000} + V_{00i}$

Level-1 Variance of $\mathbf{e_{tdi}} \rightarrow \boldsymbol{\sigma_e^2 = 6.1046}$
Level-2 Variance of $\mathbf{U_{0di}} \rightarrow \boldsymbol{\tau_{U_0}^2 = 0}$
Level-3 Variance of $\mathbf{V_{00i}} \rightarrow \boldsymbol{\tau_{V_{00}}^2 = 6.7127}$

ICC_{L3b} for the correlation of occasions from same person:

$$ICC = \frac{6.7127}{6.7127 + 6.1046} = .52$$

Level 1 Time: $\text{Fatigue}_{tdi} = \beta_{0di} + e_{tdi}$

Level 2 Day: Intercept: $\beta_{0di} = \delta_{00i} + U_{0di}$

Level 3 Person: Intercept: $\delta_{00i} = \gamma_{000} + V_{00i}$

Proportion variance at each level:	$\mathbf{ICC_{L3b}}$ **for proportion of between-person variance over total**
Total = 4.5427 + 1.6826 + 6.4470 = 12.672	**variance** = 6.4470 / 12.672 = .51
Level 1 (time) = 4.5427 / 12.672 = .36 Level 2 (day) = 1.6826 / 12.672 = .13 Level 3 (person) = 6.4470 / 12.672 = .51	$\mathbf{ICC_{L2b}}$ **for proportion of between-day over within-person variance** = 1.6826 / (1.6826 + 4.5427) = .27 → This ICC_{L2b} = .27 is significantly greater than 0 via −2ΔLL for 3- vs. 2-level.

Within-Day Negative Affect: 2 levels or 3?

Level 1 Time: $\text{NegAff}_{tdi} = \beta_{0di} + e_{tdi}$

Level 2 Day: Intercept: $\beta_{0di} = \delta_{00i} + 0$

Level 3 Person: Intercept: $\delta_{00i} = \gamma_{000} + V_{00i}$

Level-1 Variance of $\mathbf{e_{tdi}}$ → $\boldsymbol{\sigma^2_e = 5.88}$
Level-2 Variance of $\mathbf{U_{0di}}$ → $\boldsymbol{\tau^2_{U_0} = 0}$
Level-3 Variance of $\mathbf{V_{00i}}$ → $\boldsymbol{\tau^2_{V_{00}} = 19.60}$

ICC_{L3b} for the correlation of occasions from same person:

$$\text{ICC} = \frac{19.599}{19.599 + 5.865} = .77$$

Level 1 Time: $\text{NegAff}_{tdi} = \beta_{0di} + e_{tdi}$

Level 2 Day: Intercept: $\beta_{0di} = \delta_{00i} + U_{0di}$

Level 3 Person: Intercept: $\delta_{00i} = \gamma_{000} + V_{00i}$

Proportion variance at each level:	**ICC_{L3b} for proportion of between-person variance over total variance** = 19.353 / 25.263 = .77
Total = 5.179 + 0.731 + 19.353 = 25.263	
Level 1 (time) = 5.179 / 25.263 = .20 Level 2 (day) = 0.731 / 25.263 = .03 Level 3 (person) = 19.353 / 25.263 = .77	**ICC_{L2b} for proportion of between-day over within-person variance** = 0.731 / (0.731 + 5.179) = .12 → This ICC_{L2b} = .12 is significantly greater than 0 via −2ΔLL for 3- vs. 2-level.

Previous Nights' Sleep: 1 level or 2?

$$\text{Level 1 Time:}\quad \text{SleepHours}_{tdi} = \beta_{0di} + 0$$
$$\text{Level 2 Day:}\quad \text{Intercept:}\quad \beta_{0di} = \delta_{00i} + U_{0di}$$
$$\text{Level 3 Person: Intercept:}\quad \delta_{00i} = \gamma_{000} + V_{00i}$$

ICC_{L3b} for the correlation of days from same person:

$$ICC = \frac{0.5421}{0.5421 + 1.4686} = .269$$

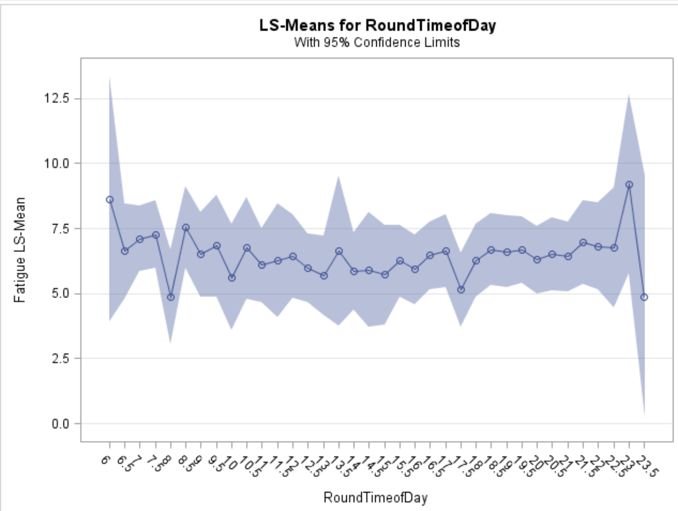
Time of Day: 2 levels or 3?

$$\text{Level 1 Time:}\quad \text{NegAff}_{tdi} = \beta_{0di} + e_{tdi}$$
$$\text{Level 2 Day:}\quad \text{Intercept:}\quad \beta_{0di} = \delta_{00i} + U_{0di}$$
$$\text{Level 3 Person: Intercept:}\quad \delta_{00i} = \gamma_{000} + V_{00i}$$

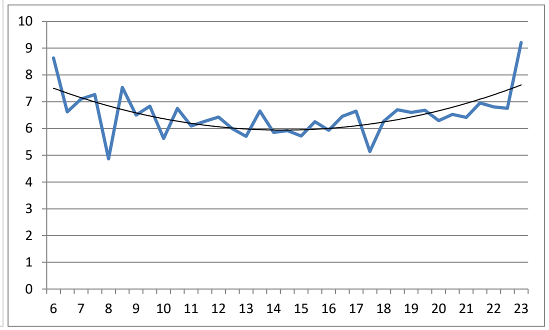
Proportion variance at each level: Total = 23.793 + 0 + 0.580 = 24.373 Level 1 (time) = 23.793 / 24.373 = .98 Level 2 (day) = 0 / 24.373 = 0 Level 3 (person) = 0.580 / 24.373 = .02s	**ICC_{L3b} for proportion of between-person variance over total variance =** 0.5798 / 24.373 = .02 **ICC_{L2b} for proportion of between-day over within-person variance** = 0

Examining Change over *Time*

Level 1: Within-day change in fatigue (rounded to 30 min)

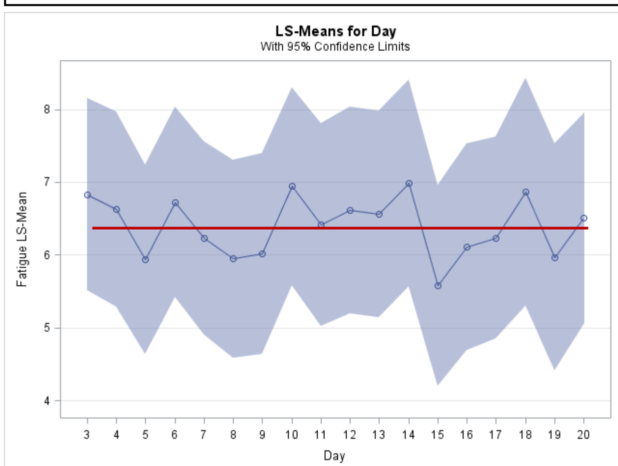


It looks like fatigue has a quadratic pattern of change over the day



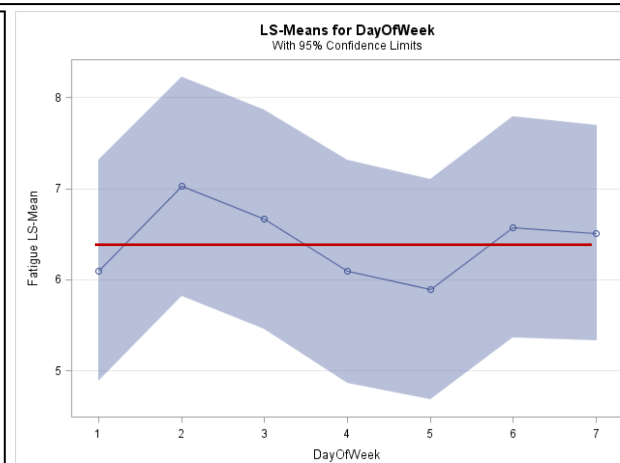
Even though this is not a study about within-person change per se, it's still possible for fixed and random effects of *time* to be relevant… which need to be addressed before examining effects of other predictors!

Level 2: Across-day change in Fatigue: Two possibilities



No systematic change by day of study (3–20)

But Day 2 is slightly higher, so it looks like these kids may have a case of the Mondays! 😊



Fatigue Unconditional Models for Time

Level 1 Time: $\text{Fatigue}_{tdi} = \beta_{0di} + \beta_{1di}(\text{Hour}_{tdi} - 15) + \beta_{2di}(\text{Hour}_{tdi} - 15)^2 + e_{tdi}$

Level 2 Day: Intercept: $\beta_{0di} = \delta_{00i} + \delta_{01i}(\text{Monday}_{di}) + U_{0di}$

Linear Time: $\beta_{1di} = \delta_{10i} + U_{1di}$

Quadratic Time: $\beta_{2di} = \delta_{20i} + U_{2di}$

Level 3 Person: Intercept: $\delta_{00i} = \gamma_{000} + \gamma_{001}(\overline{\text{Hour}_i} - 15) + V_{00i}$

Linear Time: $\delta_{10i} = \gamma_{100} + \gamma_{101}(\overline{\text{Hour}_i} - 15) + V_{10i}$

Quadratic Time: $\delta_{20i} = \gamma_{200} + \gamma_{201}(\overline{\text{Hour}_i} - 15) + V_{20i}$

It's Monday: $\delta_{01i} = \gamma_{010}$

The new fixed effects include a L2 dummy code for if it's Monday, as well as linear and quadratic L1 hour of day (predictors where 0=3 PM). The linear L3 time effects (also 0=3PM) are needed to create contextual effects given the use of constant-centering for L1 time. Btw, quadratic L3 time main and interaction effects were tested and nonsignificant.

- **Person-level-3 sources of variance to be predicted:**
  - Intercept V_{00i}: Why do some kids report more fatigue at 3 PM than other kids?
  - Linear and quadratic time-of-day slopes V_{10i} and V_{20i}: Why do some kids change more in fatigue throughout the day than other kids?
  - (Monday slope V_{01i} was NS: Why do some kids report more fatigue when it's Monday than other kids?)
  - *Scale-model random intercept not included: Why are some kids more inconsistent in fatigue than others?*
- **Day-level-2 sources of variance to be predicted:**
  - Intercept U_{0di}: Why is more fatigue at 3 PM reported on some days than other days?
  - Linear and quadratic time-of-day slopes V_{10i} and V_{20i}: Why is more change in fatigue throughout the day reported on some days than other days?
- **Time-level-1 sources of variance to be predicted:**
  - Residual e_{tdi}: Why is fatigue higher than predicted (by everything else) on some occasions than others?

Predicting Fatigue by Negative Affect

Level 1 Time: Fatigue$_{tdi} = \beta_{0di} + \beta_{1di}(\text{Hour}_{tdi} - 15) + \beta_{2di}(\text{Hour}_{tdi} - 15)^2 + \beta_{3di}(\text{NA}_{tdi} - \overline{\text{NA}}_{di}) + e_{tdi}$

Level 2 Day: Intercept: $\beta_{0di} = \delta_{00i} + \delta_{01i}(\text{Monday}_{di}) + \delta_{02i}(\overline{\text{NA}}_{di} - \overline{\text{NA}}_{i}) + U_{0di}$

Linear Time: $\beta_{1di} = \delta_{10i} + U_{1di}$

Quadratic Time: $\beta_{2di} = \delta_{20i} + U_{2di}$

Within-Day NA: $\beta_{3di} = \delta_{30i} + U_{3di}$ *Also random slope across days*

Level 3 Person: Intercept: $\delta_{00i} = \gamma_{000} + \gamma_{001}(\overline{\text{Hour}}_{i} - 15) + \gamma_{002}(\overline{\text{NA}}_{i} - 10) + V_{00i}$

Linear Time: $\delta_{10i} = \gamma_{100} + \gamma_{101}(\overline{\text{Hour}}_{i} - 15)$

Quadratic Time: $\delta_{20i} = \gamma_{200} + \gamma_{201}(\overline{\text{Hour}}_{i} - 15) + V_{20i}$

Within-Day NA: $\delta_{30i} = \gamma_{300}$

It's Monday: $\delta_{01i} = \gamma_{010}$

Between-Day NA: $\delta_{02i} = \gamma_{020}$

Given their centering, **each NA effect is specific to its level**—*being grumpier than*:

L1 = the rest of the day (the day's mean)

L2 = usual (kid's mean across days)

L3 = other people (on average, 0 = 10)

	Effect	Estimate	Std Err	DF	t Value	Pr > |t|
γ_{000}	Intercept	7.2351	0.3269	27.5	22.13	<.0001
γ_{010}	Monday	0.4678	0.2691	243	1.74	0.0834
γ_{010}	L1T	0.008782	0.01598	254	0.55	0.5831
γ_{010}	L1T*L1T	0.01153	0.006086	14.3	1.89	0.0785
γ_{001}	L3T	-0.1144	0.2599	21.6	-0.44	0.6640
γ_{101}	L1T*L3T	-0.00765	0.01498	253	-0.51	0.6101
γ_{201}	L1T*L1T*L3T	-0.00890	0.005504	15.6	-1.62	0.1259
γ_{300}	L1NA	0.2283	0.03145	631	7.26	<.0001
γ_{020}	L2NA	0.4293	0.06821	227	6.29	<.0001
γ_{020}	L3NA	0.6205	0.08167	27.1	7.60	<.0001

Predicting Fatigue by Hours of Sleep

$$\text{Level 1 Time:} \quad \text{Fatigue}_{tdi} = \beta_{0di} + \beta_{1di}(\text{Hour}_{tdi} - 15) + \beta_{2di}(\text{Hour}_{tdi} - 15)^2 + \beta_{3di}(\text{NA}_{tdi} - \overline{\text{NA}}_{di}) + e_{tdi}$$

$$\text{Level 2 Day:} \quad \text{Intercept:} \quad \beta_{0di} = \delta_{00i} + \delta_{01i}(\text{Monday}_{di}) + \delta_{02i}(\overline{\text{NA}}_{di} - \overline{\text{NA}}_{i}) + \delta_{03i}(\text{Sleep}_{di} - \overline{\text{Sleep}}_{i}) + U_{0di}$$

$$\text{Linear Time:} \quad \beta_{1di} = \delta_{10i} + U_{1di}$$

$$\text{Quadratic Time:} \quad \beta_{2di} = \delta_{20i} + U_{2di}$$

$$\text{Within-Day NA:} \quad \beta_{3di} = \delta_{30i} + U_{3di}$$

$$\text{Level 3 Person: Intercept:} \quad \delta_{00i} = \gamma_{000} + \gamma_{001}(\overline{\text{Hour}}_{i} - 15) + \gamma_{002}(\overline{\text{NA}}_{i} - 10) + \gamma_{003}(\overline{\text{Sleep}}_{i} - 7) + V_{00i}$$

$$\text{Linear Time:} \quad \delta_{10i} = \gamma_{100} + \gamma_{101}(\overline{\text{Hour}}_{i} - 15)$$

$$\text{Quadratic Time:} \quad \delta_{20i} = \gamma_{200} + \gamma_{201}(\overline{\text{Hour}}_{i} - 15) + V_{20i}$$

$$\text{Within-Day NA:} \quad \delta_{30i} = \gamma_{300}$$

$$\text{It's Monday:} \quad \delta_{01i} = \gamma_{010}$$

$$\text{Between-Day NA:} \quad \delta_{02i} = \gamma_{020}$$

$$\text{Yesterday Sleep:} \quad \delta_{03i} = \gamma_{030} + V_{03i}$$

Tried random slope across persons, but did not converge

Given their centering, **each sleep effect is specific to its level**—*less sleep than*:

L1 = none
L2 = usual (kid's mean across days)
L3 = other kids (0 = 7 hours)

	Effect	Estimate	Std Err	DF	t Value	Pr > |t|
(effects related to time omitted for brevity)						
γ_{300}	L1NA	0.2283	0.03145	631	7.26	<.0001
γ_{020}	L2NA	0.4293	0.06821	227	6.29	<.0001
γ_{020}	L3NA	0.6205	0.08167	27.1	7.60	<.0001
γ_{030}	L2Sleep	-0.1472	0.08092	208	-1.82	0.0703
γ_{003}	L3Sleep	-0.2308	0.3056	30.5	-0.76	0.4559

Getting less sleep than usual the night before is (almost) related to feeling more fatigue that day. But getting less sleep than other people does *not* imply that you report feeling more fatigue than other kids.

Sleep Moderating Effects of Negative Affect

$$\text{Level 1 Time: } \quad \text{Fatigue}_{tdi} = \beta_{0di} + \beta_{1di}(\text{Hour}_{tdi} - 15) + \beta_{2di}(\text{Hour}_{tdi} - 15)^2 + \beta_{3di}(\text{NA}_{tdi} - \overline{\text{NA}}_{di}) + e_{tdi}$$

$$\text{Level 2 Day: } \quad \text{Intercept: } \beta_{0di} = \delta_{00i} + \delta_{01i}(\text{Monday}_{di}) + \delta_{02i}(\overline{\text{NA}}_{di} - \overline{\text{NA}}_{i}) + \delta_{03i}(\text{Sleep}_{di} - \overline{\text{Sleep}}_{i}) + \delta_{04i}(\overline{\text{NA}}_{di} - \overline{\text{NA}}_{i})(\text{Sleep}_{di} - \overline{\text{Sleep}}_{i}) + U_{0di}$$

$$\text{Linear Time: } \beta_{1di} = \delta_{10i} + U_{1di}$$

$$\text{Quadratic Time: } \beta_{2di} = \delta_{20i} + U_{2di}$$

$$\text{Within-Day NA: } \beta_{3di} = \delta_{30i} + \delta_{33i}(\text{Sleep}_{di} - \overline{\text{Sleep}}_{i}) + U_{3di}$$

$$\text{Level 3 Person: Intercept: } \delta_{00i} = \gamma_{000} + \gamma_{001}(\overline{\text{Hour}}_{i} - 15) + \gamma_{002}(\overline{\text{NA}}_{i} - 10) + \gamma_{003}(\overline{\text{Sleep}}_{i} - 7) + \gamma_{004}(\overline{\text{NA}}_{i} - 10)(\overline{\text{Sleep}}_{i} - 7) + V_{00i}$$

$$\text{Linear Time: } \delta_{10i} = \gamma_{100} + \gamma_{101}(\overline{\text{Hour}}_{i} - 15)$$

$$\text{Quadratic Time: } \delta_{20i} = \gamma_{200} + \gamma_{201}(\overline{\text{Hour}}_{i} - 15) + V_{20i}$$

$$\text{Within-Day NA: } \delta_{30i} = \gamma_{300} + \gamma_{303}(\overline{\text{Sleep}}_{i} - 7)$$

$$\text{It's Monday: } \delta_{01i} = \gamma_{010}$$

$$\text{Between-Day NA: } \delta_{02i} = \gamma_{020} + \gamma_{023}(\overline{\text{Sleep}}_{i} - 7)$$

$$\text{Yesterday Sleep: } \delta_{03i} = \gamma_{030} + \gamma_{002}(\overline{\text{NA}}_{i} - 10)$$

$$\text{Between-Day NA by Yesterday Sleep: } \delta_{04i} = \gamma_{040}$$

$$\text{Within-Day NA by Yesterday Sleep: } \delta_{33i} = \gamma_{330}$$

Does the relationship between fatigue and negative affect differ by previous hours of sleep?

To answer this question, I added **6 interactions** (of L2 and L3 sleep with each level of NA).

Big-picture interpretation of cross-level interaction results:
Kids who sleep more than others (level 3) have *weaker* within-day (level-1) and between-day (level-2) effects of negative affect on fatigue—they appear *less* susceptible to feeling more tired when they are grumpy (or vice-versa).

But kids who sleep more than others (level 3) have a *greater* BP effect of negative affect—the tendency for grumpy kids to be tired kids is stronger in kids who sleep more.

Three-Level Longitudinal Models for Ecological Momentary Assessment Data

- Topics:
  - Example 6: Predicting fatigue in EMA data
  - **More general info about three-level models**

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

- **Variable-Centering**: Omitting a fixed effect assumes that the slope at that level <u>does not exist</u> (= 0)
  - Remove L3 effect? Assume L3 Between-Person effect = 0
    - *L1 effect = Within-Day effect, L2 effect = Between-Day effect*
  - Then remove L2 effect? Assume L2 Between-Day effect = 0
    - *L1 effect = Within-Day effect*

- **Constant-Centering**: Omitting a fixed effect means the slope at that level is <u>equivalent to</u> the slope at the level below
  - Remove L3 effect? Assume L3 Between-Person = L2 Between-Day effect
    - *L1 effect = Within-Day effect, L2 effect = 'smushed' BP and BD effects*
  - Then remove L2 effect? Assume L2 Between-Day effect = L1 effect
    - *L1 'smushed' = Within-Day, Between-Day, and Between-Person effects*

Summary: Three-Level Random Effects Models

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

Pseudo-R^2 in Three-Level Models

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