EPSY 8268 Assignment 5

To complete this assignment, identify a data set with three levels, with sufficient numbers of units at each level – hopefully at least 15 cases within group at level 1, 10 units within group at level 2 and 15 level-3 groups. Smaller samples at each level will work for the assignment, but you might have trouble with estimation and power (statistical significance). The data for this assignment can include longitudinal data, where time points are at level 1. Remember, there are lots of data sets online – there are a few links to data at the class website.

You may use the same data set from Assignment 4 (based on the *US Sustaining Effects Study*, an example data set used in the HLM Users’ Guide and provided in the example data accompanying the HLM software). A different sample is included here, summarized below.

***Growth in Mathematics Study - Modified***

A longitudinal study of children’s growth in mathematics achievement during primary school, from grades 1 to 5. This study focuses on children with multiple at-risk indicators. Level-1 provides an opportunity to model growth curves per student, Level-2 includes student characteristics, and Level-3 includes school characteristics.

Level-1 Data: EG1r.sav

 7230 observations for 1700 children, beginning in grade 1, annually through grade 5.

* Year (year of the study minus 3.5, 0 = the midpoint of the study)
* Grade (the grade of the student minus 1, so that 0 = grade 1)
* Math (the math score on an IRT scale score metric; *M* = -0.5, *SD* = 1.5)
* Retained (1 = retained in grade during that year, 0 = not retained; 5% time points)

Level-2 Data: EG2r.sav

 1700 children, with full information (no missing data at level-2).

* Female (1 = female, 0 = male; 51% female)
* Black (1 = African American, 0 = Not AA; 69% AA)
* Latino (1 = Latino, 0 = not Latino; 15% Latino)
* SOC (1 = Student of Color, 0 = White; 84% SOC)
* Retained (0 = never retained, 1 = retained at least once; 21% retained at least once)
* posid (Positive Identify, a positive outlook and belief in oneself: *M* = 50, *SD* = 10)
* ses (social-economic status, composite of several variables: *M* = 0, *SD* = 1)

Level-3 Data: EGr.sav

 60 schools with full information (no missing data at level-3).

* Size (number of students enrolled in the school; *M* = 643, *SD* = 317)
* Lowinc (percent of students from families with low incomes; *M* = 74%, *SD* = 27%)
* Mobile (percent of students that move during a school year; *M* = 35%, *SD* = 13%)
* p\_retain (percent ever retained; *M* = 21%, *SD* = 20%)
* p\_soc (percent students of color; *M* = 78%, *SD* = 27%)
* m\_ses (mean SES; *M* = .04, *SD* = .46)
* m\_posid (mean positive identity; *M* = 50.4, *SD* = 2.6)

**Assignment Tasks**

Follow the overall guidance for model building, especially when monitoring variance explained.

1. Specify the level-1 model. This will be the base model, with the time-point variable, but it may also include a time-varying covariate.
2. Specify the level-2 intercept model, followed by the slope(s) model.
3. Specify the level-3 intercept model, followed by the slope(s) model.

Be sure to include group means in the intercept model for both levels 2 and 3 when there are explanatory variables in prior models, including any main effects that need to be accounted for when including explanatory variables in slopes models (interactions).

Report the following:

1. Estimate an unconditional model or a baseline model (if growth model).
	1. Briefly describe the data set being used, the variables included in the model (such as reported above) for the assignment and report the software being used.
	2. Estimate the unconditional/baseline model. Report a table of the fixed and random effects, interpreting each coefficient. Do this in a list format.
	3. Report a 95% confidence interval around the fixed effect(s).
	4. Report the range of plausible values for the fixed effect(s) (the range of 95% of the intercepts and slopes).
	5. What is the reliability of the intercept?

If you estimated a growth model:

* 1. What is the reliability of the slopes?
	2. What is the correlation between the intercepts and slopes? What is your interpretation of this correlation – what does it mean?
1. Specify a model with at least one explanatory variable at each level, including at least one explanatory variable explaining a slope at level 2 or 3 (if a growth-model, include a potential time-varying covariate; consider using ‘retention’ in the EG data, or no variables at the level-1 model). If you include a time-varying covariate at level 1, fix the effect at level 2. Clearly state whether and how each variables is centered.
	1. Write out the model in hierarchical notation.
	2. Report a table of the fixed and random effects as each level is developed (to keep track of baseline random effects variance components).
	3. Define and interpret the meaning of each fixed effect and random variance component. [Again, create a list for this.]
	4. Report the variance explained for each model at each level with explanatory variables. Be sure to attend the which variance is the appropriate one regarding baseline.
	5. Provide a general statement about the effects of the variables included in the model: What is the overall conclusion regarding the outcome as measured and modeled? Make a brief statement about the role of each level.