General MLMs for Two-Level Cross-Classified Data

- Topics:
  - Cross-sectional cross-classification (time-invariant groups)
  - Longitudinal cross-classification (time-varying groups)
  - Random slopes and smushing in cross-classified models

More Complex Multilevel Designs

- Multilevel models are specified based on the relevant dimensions by which observations differ each other, and how the units are organized

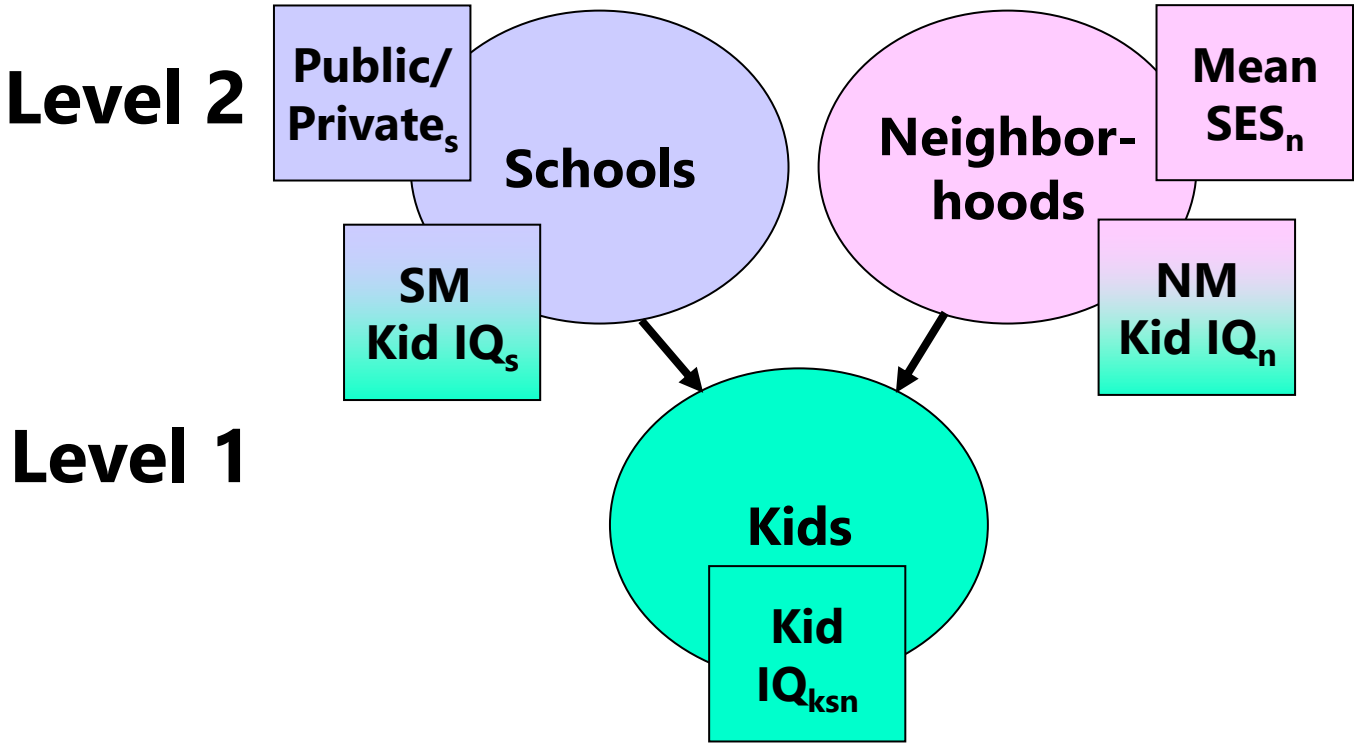
- Two-level models have at least two piles of variance, in which level-1 units are **nested** within level-2 units:
  - Longitudinal data: Occasions nested within Persons
  - Clustered data: Students nested within Teachers

- Three-level models have at least three piles of variance, in which level-2 units are **nested** within level-3 units (stay tuned):
  - Longitudinal data: Occasions nested within Persons within Families
  - Clustered data: Students nested within Teachers within Schools

- In other designs, multiple sources of systematic variation may be present, but the sampling may be **crossed** instead…
  - Models with crossed random effects are known as "**cross-classified**" (if 1 for each) or "**multiple membership**" (if weights sum to 1) models in clustered data
  - Here is a more extended treatment by Don Hedeker than what I have time to do
  - Here are a few examples on when this might happen…

Kids, Schools, and Neighborhoods

- Kids are nested within schools AND within neighborhoods
- Not all kids from same neighborhood live in same school, so schools and neighborhoods are crossed dimensions at level 2
- Can include predictors for each source of variation



Specifying Cross-Classified Models

- If there is **only one L1 observation** per combination of L2 crossed units, then their interaction = residual variance
  - e.g., Only one trial per combination of subject and item? Then:
  - L2 subject random intercept = subject mean differences
  - L2 item random intercept = item mean differences
  - L1 residual = subject by item interaction

- If there is **more than one L1 observation** per combo of L2 crossed units, their interaction can have a L2 random intercept
  - e.g., 2+ kids from same school and neighborhood? Then:
  - L2 school random intercept = school mean differences
  - L2 neighborhood random intercept = neighborhood mean differences
  - L2 school by neighborhood random intercept = school by neighborhood interaction (creates extra correlation within crossing)
  - L1 residual variance = kid-to-kid diffs within same crossing

Specifying Cross-Classified Models

- L1 predictors can have **random slopes** of over both types of L2 units, AND L2 predictors can have random slopes across the other crossed L2 dimension(s)
  - Example: L1 kids within L2 schools by L2 neighborhoods
    - L1 kid slopes could vary over L2 schools AND/OR L2 neighborhoods
    - L2 school slopes could vary over L2 neighborhoods (crossed at L2)
    - L2 neighborhood slopes could vary over L2 schools (crossed at L2)

- **Prevent smushing of L1 slopes over *all* sets of L2 units!**
  - Separate contextual effects of kid predictors for all L2 dimensions
    - e.g., After controlling for kid IQ, the mean kid IQ for your school AND the mean kid IQ in your neighborhood (and the mean kid IQ for the school*neighborhood combination) may matter incrementally
    - Use cluster-mean-centering to remove each source of L2 mean differences

What about Time-Varying Clusters?

- e.g., Students are nested within classes at each occasion…
- But if students move into different classes over time…
  - Level-1 occasions are nested within level-2 students AND within level-2 classes: Students are crossed with classes at level 2
- How to model a **time-varying classroom effect**?
  - Btw, this is the basis of so-called "value-added models"
  - Btw, the extent of same-cluster dependency could vary over time, too

- Two example options (both via fixed or random effects):
  - **"Acute" effect**: Class effect active only when students are in that class
    - e.g., class effect ← teacher bias
    - Once a student is out of the class, class effect is no longer present
  - **"Transfer" effect**: Effect is active when in class AND in the future…
    - e.g., class effect ← differential learning
    - Effect stays with the student in the future (i.e., a "layered" value-added model)

Time (t), Students (s), and Classes (c)

- Custom-built intercepts for time-varying effects of classes
  - An intercept is usually a column of 1's, but ours will be 0's and 1's to serve as switches that turn on/off class effects

				Per-Year Class ID (−99 = missing)			Intercepts for Acute Effects			Intercepts for Transfer Effects		
Student ID	Class ID	Grade	Year	Year 0 Class	Year 1 Class	Year 2 Class	Year 0 Intercept	Year 1 Intercept	Year 2 Intercept	Year 0 Effect	Year 1 Effect	Year 2 Effect
101	1	3	0	1	−99	43	1	0	0	1	0	0
101	−99	4	1	1	−99	43	0	0	0	0	0	0
101	43	5	2	1	−99	43	0	0	1	1	0	1
102	3	3	0	3	21	42	1	0	0	1	0	0
102	21	4	1	3	21	42	0	1	0	1	1	0
102	42	5	2	3	21	42	0	0	1	1	1	1

Time (t), Students (s), and Classes (c)

- Hoffman (2015) Eq. 11.3: **fixed effects model** at time *t* for student *s* in classroom *c*, for classroom as a categorical time-varying predictor:
  - Allows for control of classroom differences only….

$$\begin{aligned}\text{Effort}_{tsc} = {} & \gamma_{000} + \gamma_{100}\left(\text{Year01}_{tsc}\right) + \gamma_{200}\left(\text{Year12}_{tsc}\right) + \text{U}_{0s0} + \text{e}_{tsc} \\ & + \gamma^{0}_{001}\left(\text{Class1}_{c}\right)\left(\text{Int0}_{tsc}\right) + \gamma^{0}_{002}\left(\text{Class2}_{c}\right)\left(\text{Int0}_{tsc}\right)\cdots + \gamma^{0}_{00C}\left(\text{ClassC}_{c}\right)\left(\text{Int0}_{tsc}\right) \\ & + \gamma^{1}_{001}\left(\text{Class1}_{c}\right)\left(\text{Int1}_{tsc}\right) + \gamma^{1}_{002}\left(\text{Class2}_{c}\right)\left(\text{Int1}_{tsc}\right)\cdots + \gamma^{1}_{00C}\left(\text{ClassC}_{c}\right)\left(\text{Int1}_{tsc}\right) \\ & + \gamma^{2}_{001}\left(\text{Class1}_{c}\right)\left(\text{Int2}_{tsc}\right) + \gamma^{2}_{002}\left(\text{Class2}_{c}\right)\left(\text{Int2}_{tsc}\right)\cdots + \gamma^{2}_{00C}\left(\text{ClassC}_{c}\right)\left(\text{Int2}_{tsc}\right)\end{aligned}$$

- Hoffman (2015) Eq 11.4: classrooms as year-specific random effects crossed with students (as a random effect) at level 2:
  - Controls for and quantifies classroom variances so they can be predicted!

$$\begin{aligned}\text{Effort}_{tsc} = {} & \gamma_{000} + \gamma_{100}\left(\text{Year01}_{tsc}\right) + \gamma_{200}\left(\text{Year12}_{tsc}\right) + \text{U}_{0s0} + \text{e}_{tsc} \\ & + \text{U}^{0}_{00c}\left(\text{Int0}_{tsc}\right) + \text{U}^{1}_{00c}\left(\text{Int1}_{tsc}\right) + \text{U}^{2}_{00c}\left(\text{Int2}_{tsc}\right)\end{aligned}$$

Cross-Classified Models in Software

- Some MLM software easily allows multiple sources of crossed random effects (e.g., SPSS MIXED, SAS MIXED, R lmer)

- Other MLM software must be tricked into it via 3-level models with equality constraints (implemented in STATA MIXED)
  - Create 0/1 indicator variables for ID in smaller crossed dimension
  - Create a constant = 1 to use as level-3 ID variable; give it a random effect for each ID indicator, with equal variances and 0 covariances
  - STATA mixed uses this for the smaller crossed dimension: _all: R.ID
  - I finally figured out how to add random slopes in cross-classified models in STATA—see Example 5 (thank you, Don Hedeker, again!)
  - Appears to not allow random slopes for the tricked dimension, though

- Btw, Mplus will estimate cross-classified models using Bayes