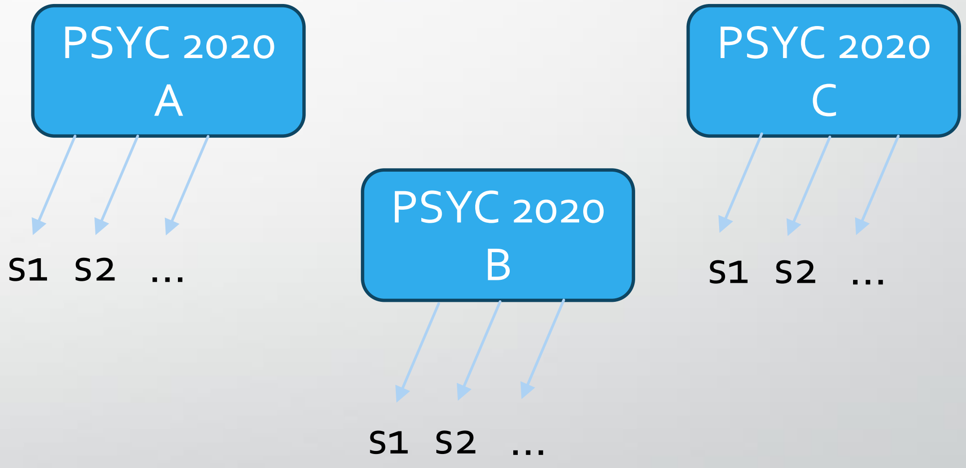
Multilevel Modeling

AKA … Hierarchical Linear Modeling, Mixed Modeling

Nesting/Clustering of Observations

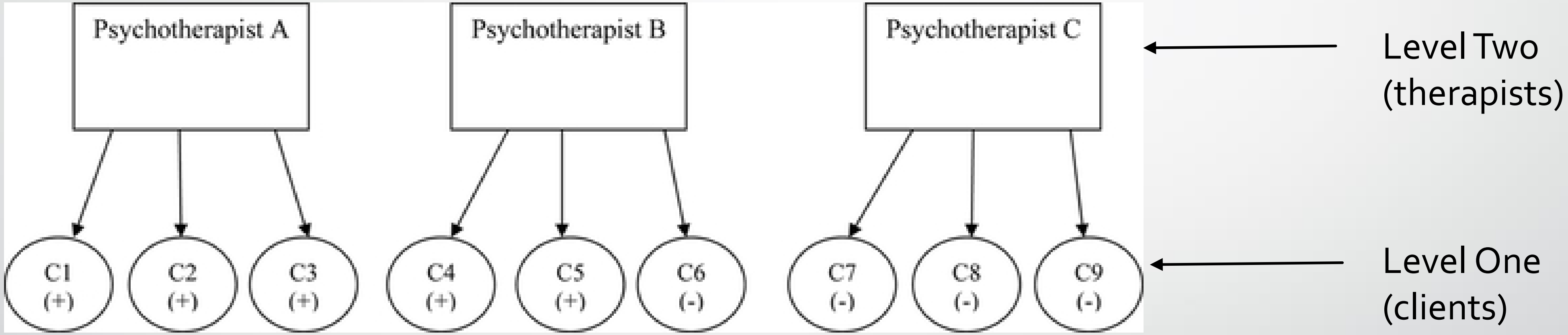
- The most famous example of nested data is *students within classes*
- Let's say we are looking at the relationship between academic self-esteem and grades; can we ignore the fact that the students are 'nested' within classes?
    - E.g., could students within the same class share some characteristics?



Some Examples of Nesting/ Clustering in Psychology

- Students within classrooms
- Students within schools
- Participants within geographic location
- Clients within therapists
- Therapists within clinics
- Clients within neighbourhoods
- Clients within clinics
- Time points within individuals
- Adolescents within peer groups

Naming the Levels



- Here, we would say that clients are *nested within therapists*
  - The lowest level, in this case *clients* (C1 … C9), is called *level one*
  - The next highest level, in this case *therapist*, is called *level two*
- Note that we can have higher orders of level (e.g., clients within therapists within clinics), but we will focus on only two-level problems for now

Why is it Important to Consider 'Nesting' or 'Clustering'?

- In the example here, clients within a given therapist appear to be more similar than clients across therapists
  - It is also possible that relationships among variables might differ within different clusterings
- What would be the consequences of ignoring which therapist treated each client?

Table 1

Simulated Data Illustrating Therapist Effects of Client Alliance Ratings

	Psychotherapist 1	Psychotherapist 2	Psychotherapist 3
Client1	6.50	—	—
Client2	5.50	—	—
Client3	6.00	—	—
Client4	—	4.50	—
Client5	—	4.00	—
Client6	—	4.00	—
Client7	—	—	4.00
Client8	—	—	3.50
Client9	—	—	3.00
Avg. Psychotherapist	6.00	4.17	3.50

Consequences of Ignoring the Clustering/Nesting

- Exaggerate the 'effective sample size'
  - Imagine that all clients with the same therapist have the same score
  - Do we really have N = 9, or do we have N =3?
- Eliminate the chance to explore within therapist vs between therapist variability, look for predictors of therapist variability, compare therapists against the average therapist, etc.

Table 1

Simulated Data Illustrating Therapist Effects of Client Alliance Ratings

	Psychotherapist 1	Psychotherapist 2	Psychotherapist 3
Client1	6.50	—	—
Client2	5.50	—	—
Client3	6.00	—	—
Client4	—	4.50	—
Client5	—	4.00	—
Client6	—	4.00	—
Client7	—	—	4.00
Client8	—	—	3.50
Client9	—	—	3.00
Avg. Psychotherapist	6.00	4.17	3.50

Consequences of Ignoring the Clustering/Nesting

- Violate the *independence* assumption
  - Clients are supposed to be independent of one another, but are 'related' because of their shared therapist

Table 1
Simulated Data Illustrating Therapist Effects of Client Alliance Ratings

	Psychotherapist 1	Psychotherapist 2	Psychotherapist 3
Client1	6.50	—	—
Client2	5.50	—	—
Client3	6.00	—	—
Client4	—	4.50	—
Client5	—	4.00	—
Client6	—	4.00	—
Client7	—	—	4.00
Client8	—	—	3.50
Client9	—	—	3.00
Avg. Psychotherapist	6.00	4.17	3.50

Introduction to the Multilevel Model (MLM)

- Recall the traditional *General Linear Model*
  - $Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \ldots + e_i \quad (i = 1, \ldots, N)$
- Fixed Coefficients
  - The coefficients (b_0, b_1, b_2, etc.) are assumed to be *fixed*, which implies that they are suitable/appropriate for (i.e., constant across) all $i = 1, \ldots, N$ units (e.g., subjects)
  - However, what if we were running a clinical study and some clients saw *Psychotherapist 1*, some saw Psychotherapist 2, etc.,
  - The model coefficients might differ across the different psychotherapists
    - E.g., b_0 may differ for subjects depending on which psychotherapist they saw

Random Coefficients

- Unlike fixed coefficients, random coefficients can take on different values
  - E.g., clients who saw *Psychotherapist 1* might have one b_0 ($b_{0,PT1}$), clients who saw *Psychotherapist 2* might have one b_0 ($b_{0,\mathrm{PT2}}$), and clients who saw Psychotherapist 3 might have one b_0 ($b_{0,PT3}$)
- We can even compute the amount of variability in the random coefficients, or compare the amount of 'between therapist' variability to the amount of 'within therapist' variability
- Before multilevel models, each of these coefficients ($b_{0,PT1}$, $b_{0,PT2}$, …) might have been calculated by analyzing the results separately for each psychotherapist
  - Multilevel models have advantages in terms of computing parameters (e.g., shrinkage/partial pooling), but we won't delve into these issues here

Between vs Within Cluster Variability

Table 1

Simulated Data Illustrating Therapist Effects of Client Alliance Ratings

	Psychotherapist 1	Psychotherapist 2	Psychotherapist 3
Client1	6.50	—	—
Client2	5.50	—	—
Client3	6.00	—	—
Client4	—	4.50	—
Client5	—	4.00	—
Client6	—	4.00	—
Client7	—	—	4.00
Client8	—	—	3.50
Client9	—	—	3.00
Avg. Psychotherapist	6.00	4.17	3.50

- Within Therapist Variability
- Between Therapist Variability
- Intraclass Correlation Coefficient
  - $ICC = \frac{Between\ Cluster\ Variability}{Between\ Cluster\ Variability + Within\ Cluster\ Variability}$

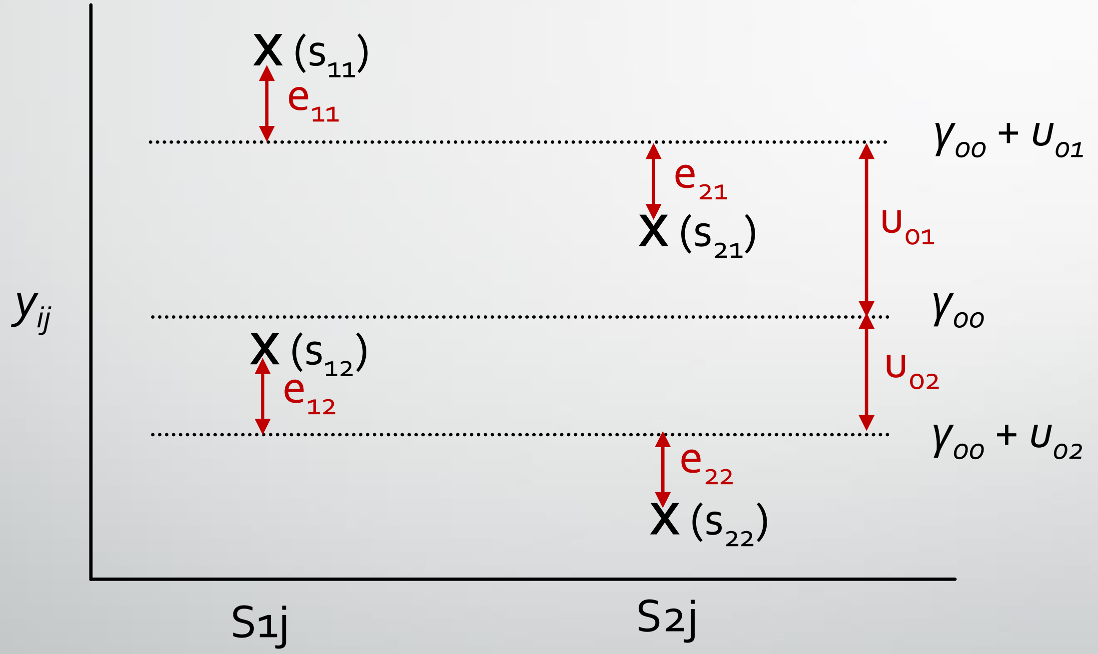
Random Intercept MLM

- In addition to computing an "average" (fixed) intercept, we are also going to allow the intercepts to vary across the level 2 units (e.g., therapists)
- Level 1 Model
  - $Y_{ij} = b_{oj} + e_{ij}$ $(i = 1, \ldots, n_j; j = 1, \ldots, J)$
  - Y_{ij} = the outcome score for the *i*th subject within the *j*th cluster
  - b_{oj} = random intercept, which can vary across the levels of *j* (level 2 variable, e.g., therapist)
  - e_{ij} = level 1 residual $(Y_{ij} - b_{oj})$, how much each unit differs from their cluster mean/intercept
- Level 2 Model
  - $b_{oj} = \gamma_{oo} + u_{oj}$
  - γ_{oo} = fixed intercept (across all *j* units)
    - The first o references its role in the level 1 model and second o references its role in the level 2 model (an intercept for an intercept …)
  - u_{oj} is a residual, indicating how much the *j*th cluster intercept deviates from the overall intercept (b_o)

Random Intercept MLM: Composite Model

- If we substitute the Level 2 model into the Level 1 model, we get the composite model:
    - Level 1: $Y_{ij} = b_{oj} + e_{ij}$, Level 2: $b_{oj} = \gamma_{oo} + u_{oj}$
  - Composite: $Y_{ij} = \gamma_{oo} + u_{oj} + e_{ij}$
- Note the two residuals, a level 1 residual (e_{ij}), and a level 2 residual (u_{oj})
  - e_{ij} represents deviations of each level 2 intercept from the observed outcome (e.g., deviations of each client from their therapist intercept, $e_{ij} = Y_{ij} - b_{oj}$)
  - u_{oj} represents deviations of each level 2 unit around the level 1 intercept (e.g., deviations of each therapist mean, b_{oj}, from the fixed intercept, γ_{oo}; $u_{oj} = b_{oj} - \gamma_{oo}$)
  - We are also interested in the variances of the residuals (σ_e^2, $\sigma_{u_0}^2$), for example to know at what levels it might be worthwhile to explore predictors

Visualizing the Random Intercept Model



- s_{11} = first unit in cluster 1
- X = data point
- No slope to any line because this is an intercept only model

Here we are just showing two 'clusters', each with two units/subjects

Adding Level 1 Predictors to the Random Intercept MLM

- We can add predictors at any level of a MLM
- Level 1 predictors need to be measured at Level 1
  - E.g., if level 1 is clients and level 2 is therapists, we need variables measured at the client level
    - E.g., client anxiety scores
- Level 1
  - $Y_{ij} = b_{0j} + b_1 X_{ij} + e_{ij}$
    - X_{ij} is level 1 variable, measured for each case *i* in each cluster *j*
    - b_1: a one unit increase in X is expected to increase Y by b_1 units
      - Note: no *j* subscript on b_1 (fixed coefficient)
  - Let's assume depression is the outcome and anxiety is a level 1 predictor
    - $DEP_{ij} = b_{0j} + b_1 ANX_{ij} + e_{ij}$
- Composite Model
  - $Y_{ij} = \gamma_{00} + u_{0j} + b_1 X_{ij} + e_{ij}$

Adding Level 2 Predictors to the Random Intercept MLM

- Level 2 predictors must be measured at Level 2 (i.e., no *i* subscript)
- Level 2
  - $b_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$
  - γ_{01}: a one unit increase in Z is expected to produce a γ_{01} unit change in b_{0j}
    - Recall (γ_{00}, γ_{01}): 1st subscript is role in the level 1 model, 2nd subscript is role in the level 2 model
- Composite Model (no level 1 predictor)
  - $Y_{ij} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} + e_{ij}$
- Back to the therapists nested within clients example:
  - $b_{0j} = \gamma_{00} + \gamma_{01}TherExp_j + u_{0j}$
    - Does the years of experience of the therapist affect the average depression level of the clients?

Random Intercept and Slope MLM

- The slope for the level one predictor is allowed to vary across the clusters (level two units)
- Level 1*:*
  - $Y_{ij} = b_{0j} + b_{1j}X_{ij} + e_{ij}$
    - Note the *j* subscript on *X*
    - b_{1j}: a one unit increase in X_{ij} is expected to increase Y_{ij} by b_{1j} units
- Level 2 (no level 2 predictors):
  - $b_{0j} = \gamma_{00} + u_{0j}$
  - $b_{1j} = \gamma_{10} + u_{1j}$
    - γ_{10} = intercept for the level 2 slope parameter
    - u_{1j} = residual for the level 2 slope parameter ($b_{1j} - \gamma_{10}$)

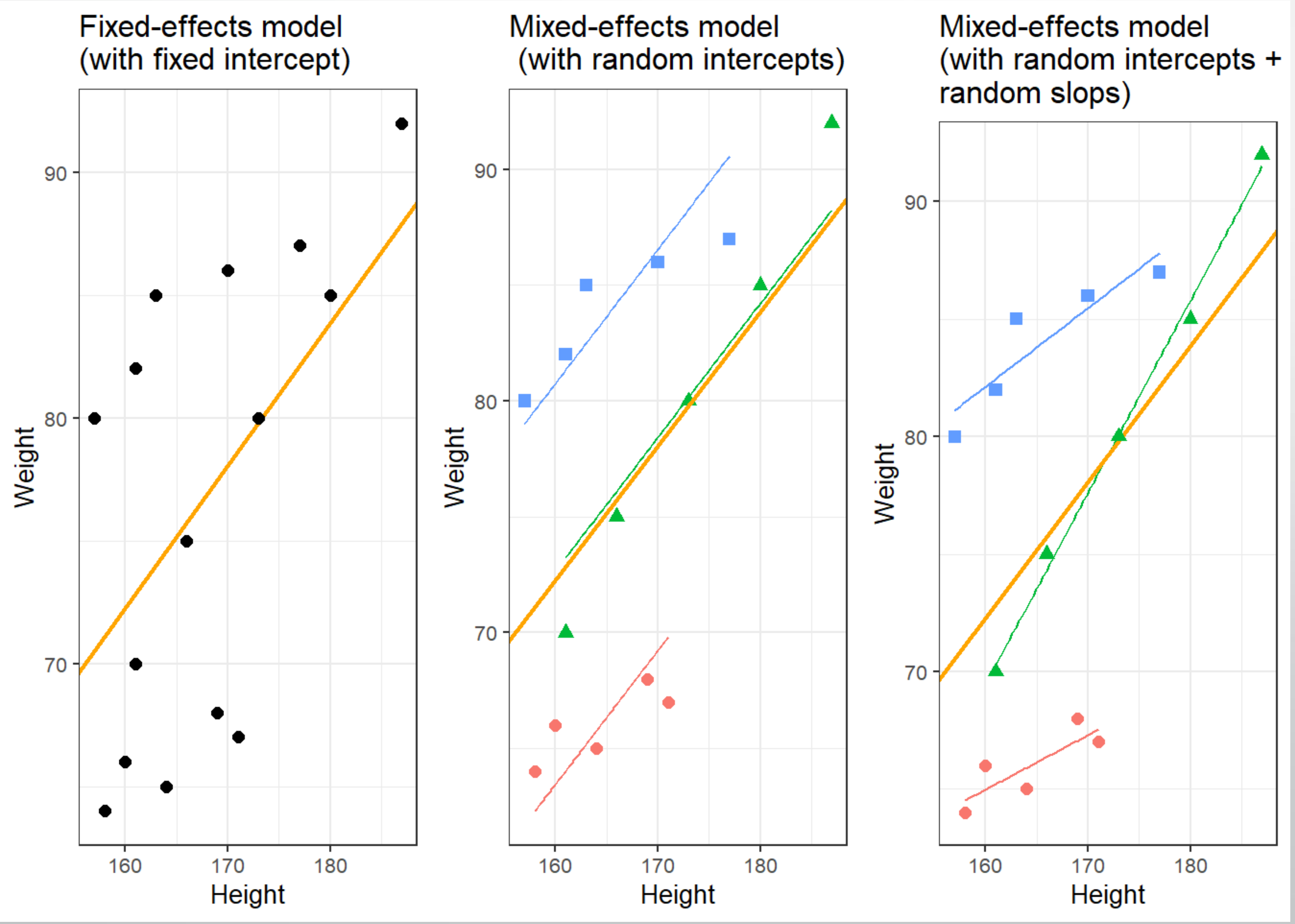
Random Intercept and Slope MLM: Composite Model

- Level 1:
  - $Y_{ij} = b_{0j} + b_{1j}X_{ij} + e_{ij}$
- Level 2 (no level 2 predictors):
  - $b_{0j} = \gamma_{00} + u_{0j}$
  - $b_{1j} = \gamma_{10} + u_{1j}$
- Composite Model
  - $Y_{ij} = \gamma_{00} + u_{0j} + (\gamma_{10} + u_{1j})\, X_{ij} + e_{ij}$
  - $Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + u_{0j} + u_{1j}X_{ij} + e_{ij}$

What Would the Data Look Like?

	ther	ther_exp	client	anx	dep
[1,]	1	8	1	17	20
[2,]	1	8	2	15	13
[3,]	1	8	3	19	23
[4,]	1	8	4	19	16
[5,]	2	6	5	18	11
[6,]	2	6	6	17	20
[7,]	2	6	7	24	25
[8,]	2	6	8	30	25
[9,]	3	13	9	19	16
[10,]	3	13	10	21	20
[11,]	3	13	11	17	22
[12,]	3	13	12	21	19

Visualizing Random Intercept/Slope Models



Excluded Topics

- We have just scraped the surface of multilevel models
- For example, we have not discussed:
  - Effect sizes
  - Assumptions
  - Comparing Models
  - Cross-Level Interactions
  - Etc., etc., etc.