Multilevel Models

Statistical Theory and Methods (HT19)

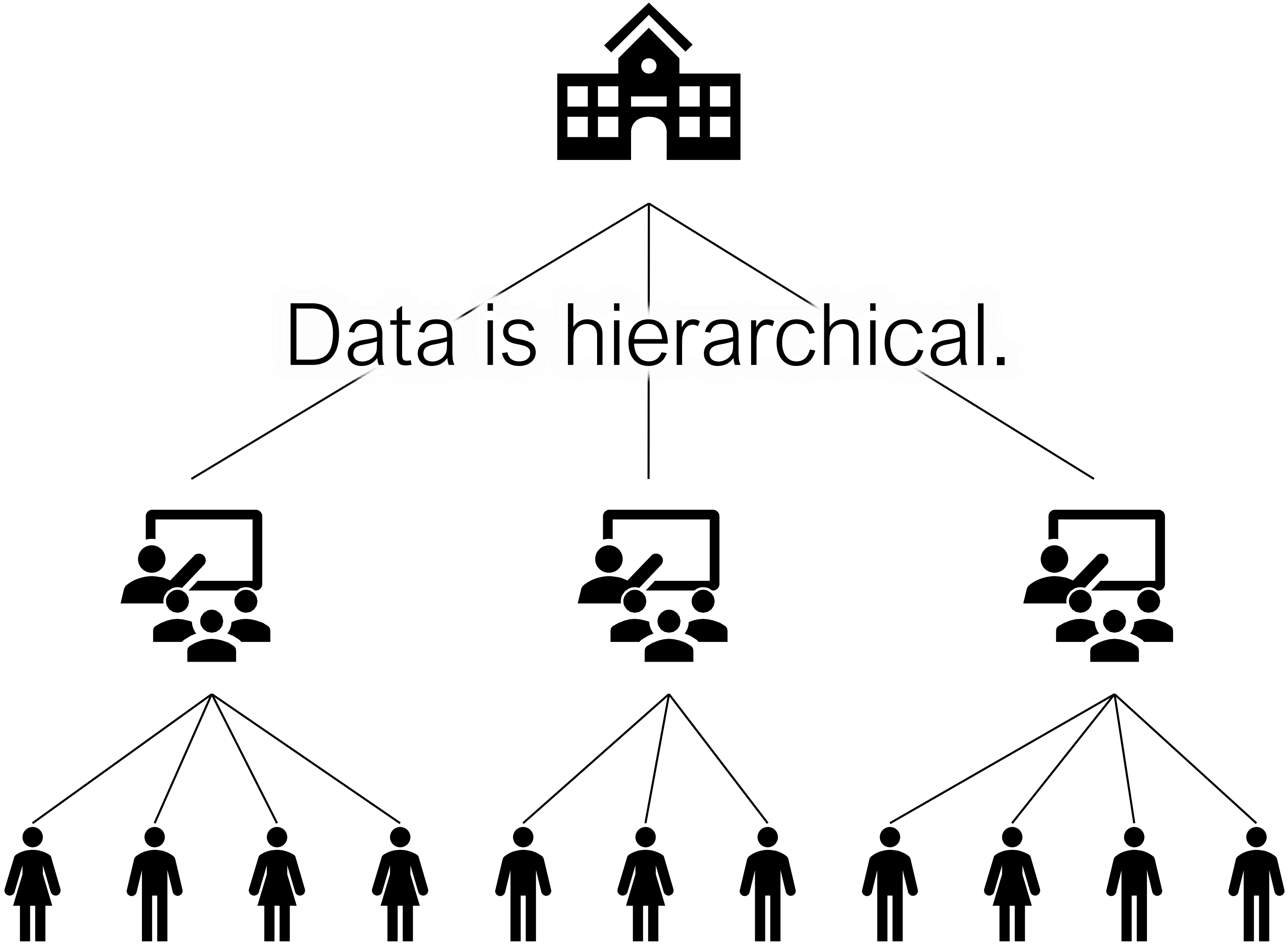
Nils Reimer

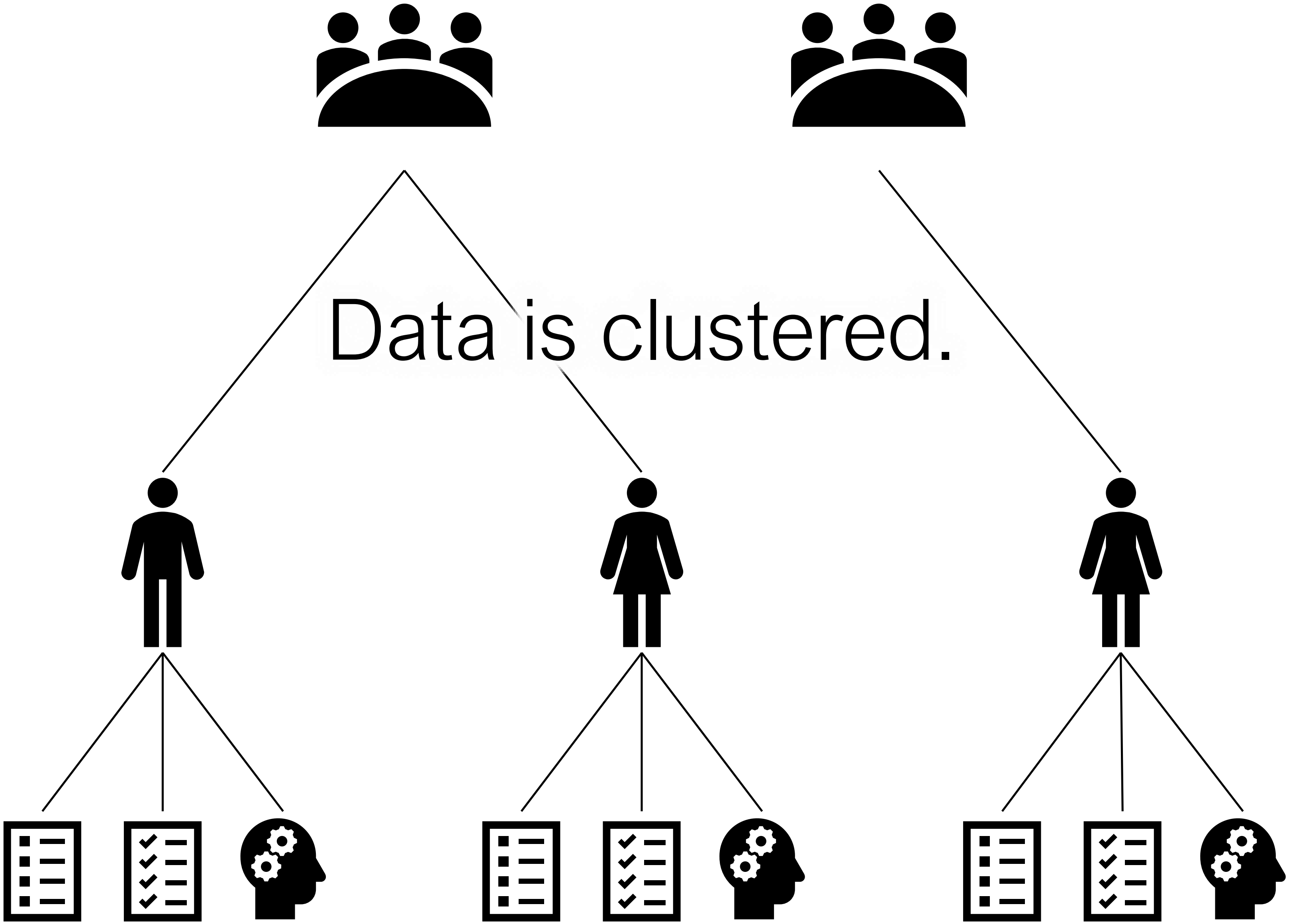
Website: nilsreimer.com Twitter: @reimthyme

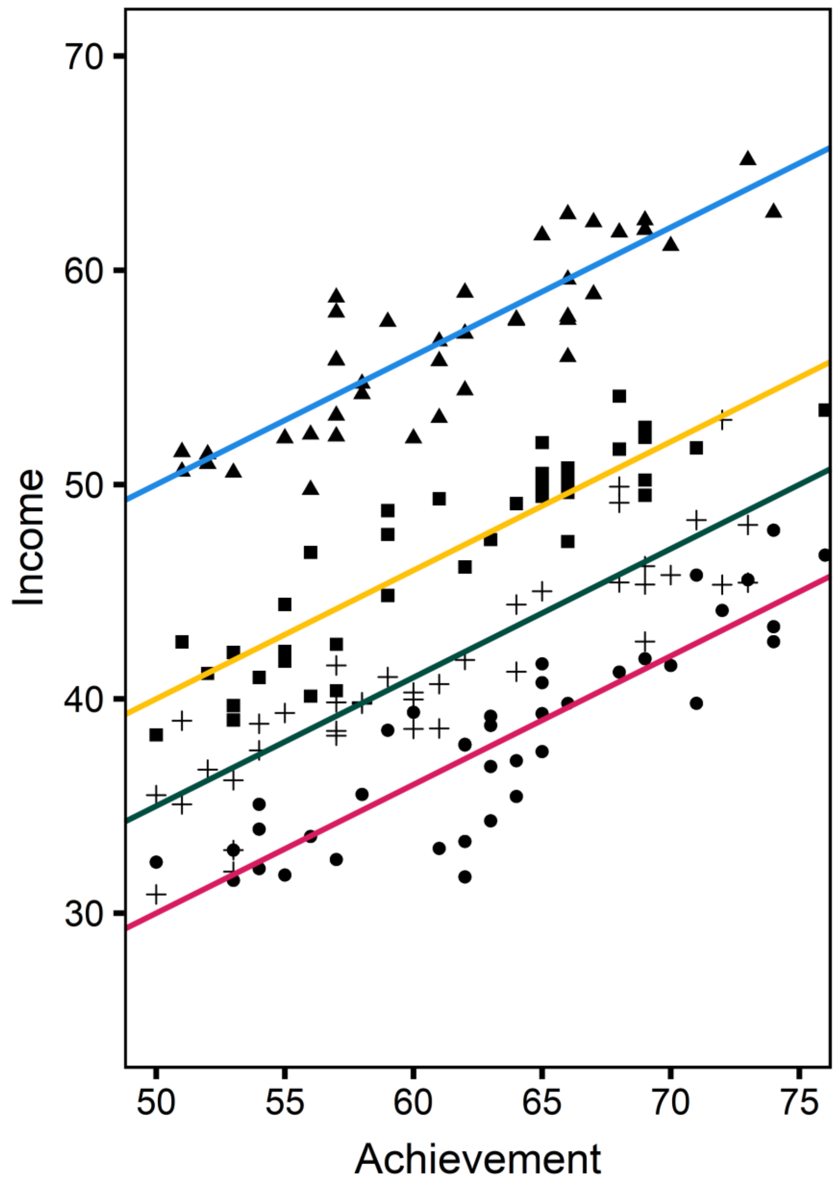
Why use multilevel models?

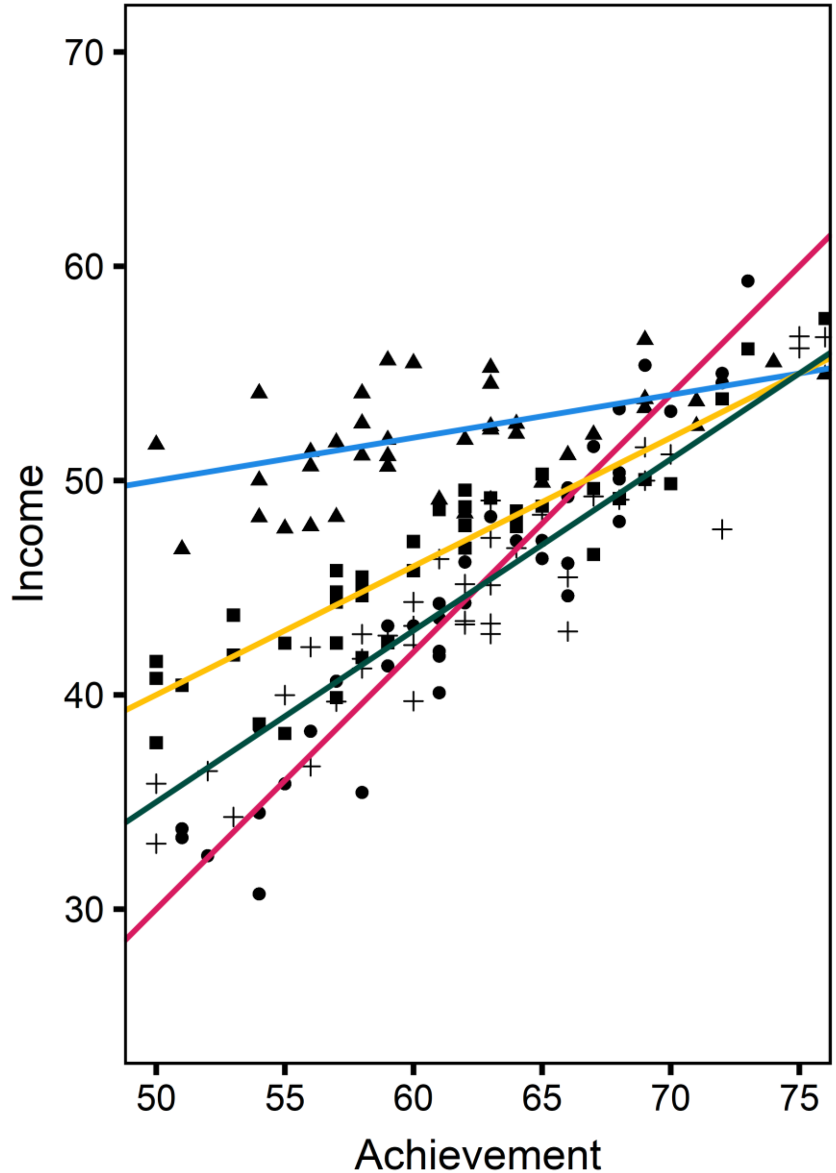
1. Data is nested.

Data is hierarchical.




Data is clustered.


70
60
50
40
30
Income
50
55
60
65
70
75
Achievement


70
60
50
40
30
Income
50
55
60
65
70
75
Achievement

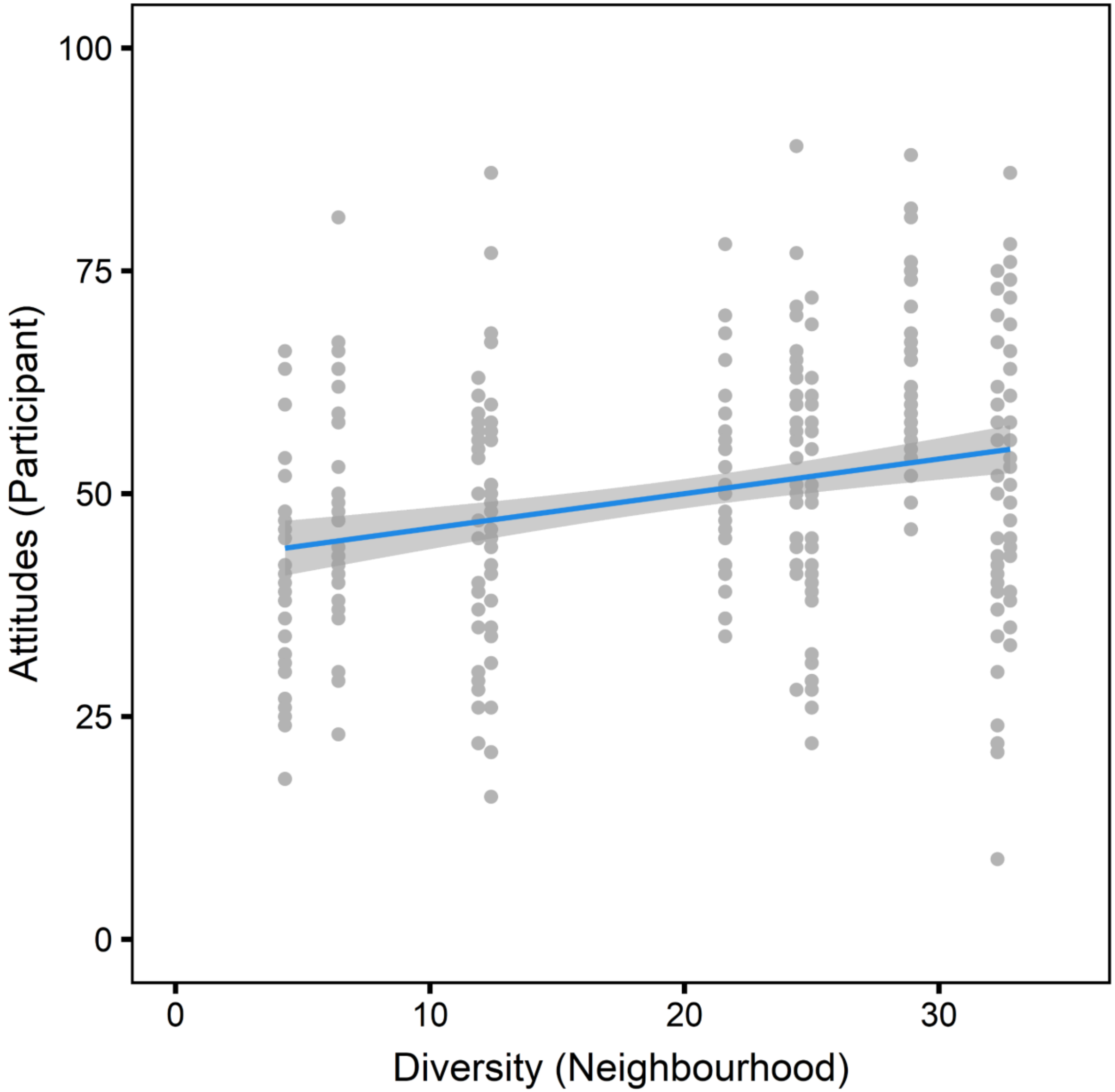
Why use multilevel models?

1. Data is nested.
   If we ignore nestedness, we misrepresent the data.

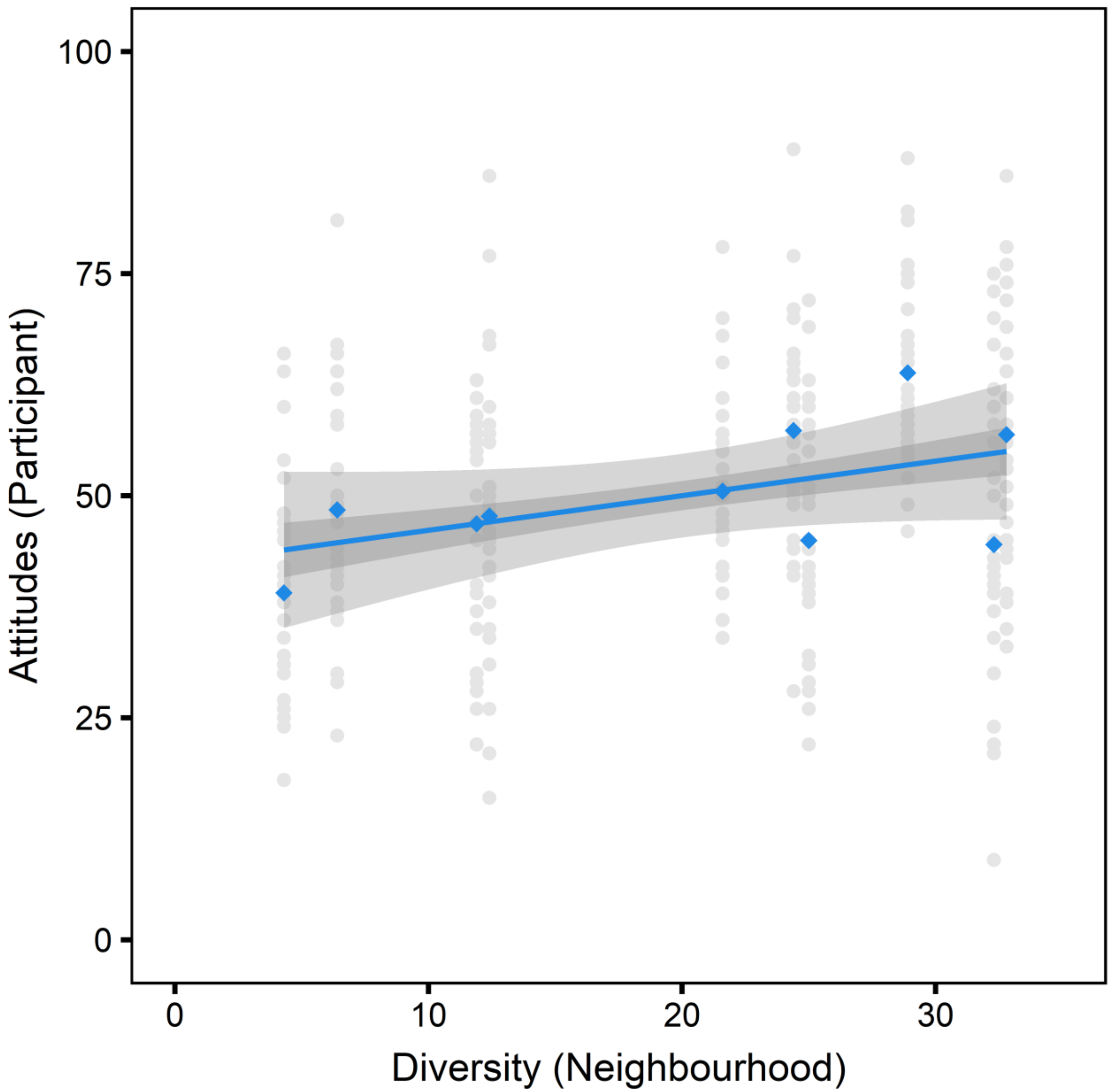
Why use multilevel models?

1. Data is nested.
   If we ignore nestedness, we misrepresent the data.

2. Ignoring nestedness does not work.



100
75
50
25
0
Attitudes (Participant)
0
10
20
30
Diversity (Neighbourhood)


Attitudes (Participant)
100
75
50
25
0
0
10
20
30
Diversity (Neighbourhood)

Why use multilevel models?

1. Data is nested.
   If we ignore nestedness, we misrepresent the data.

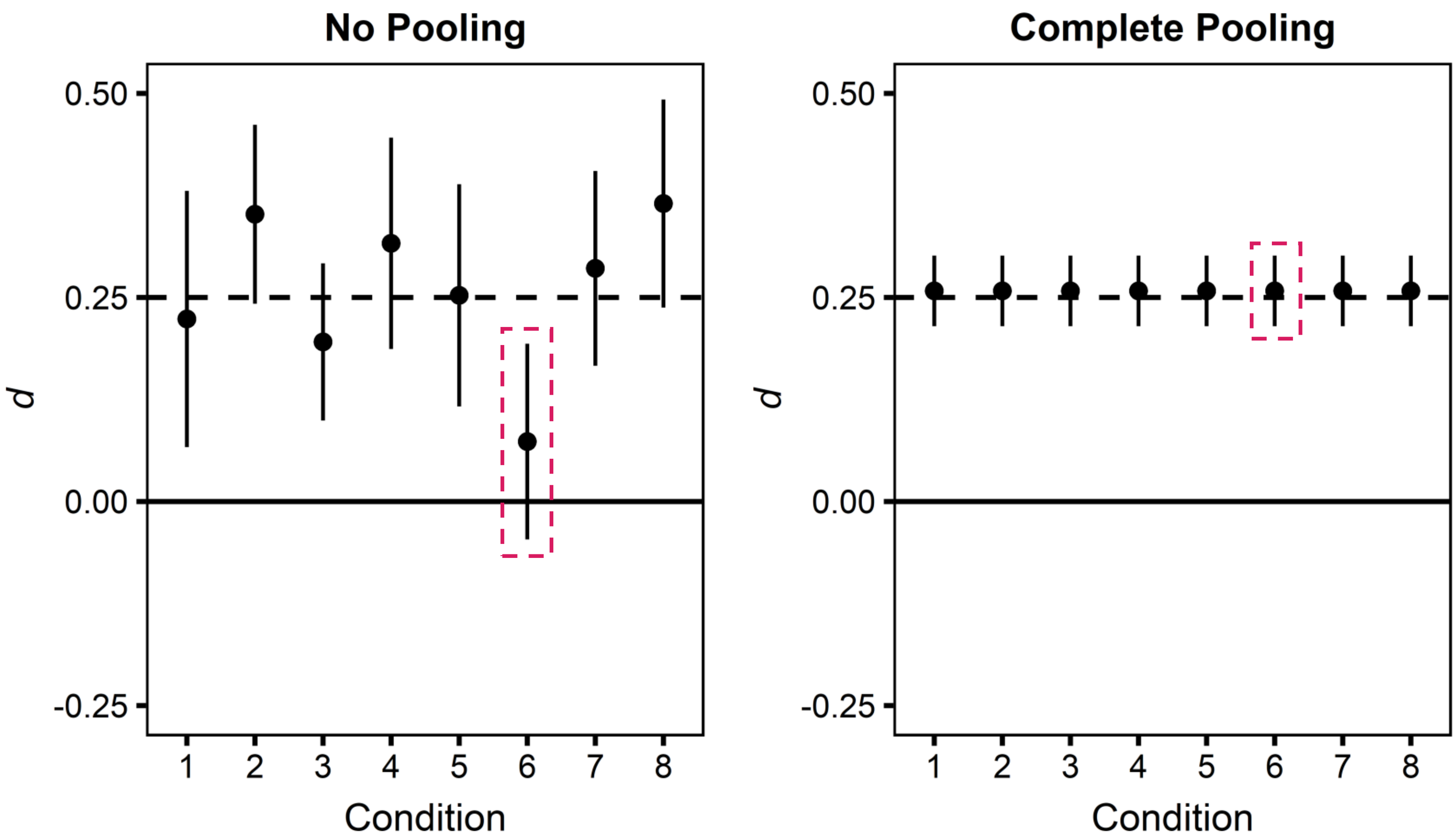
2. Ignoring nestedness does not work.
   If we pool across groups, we underestimate standard errors.

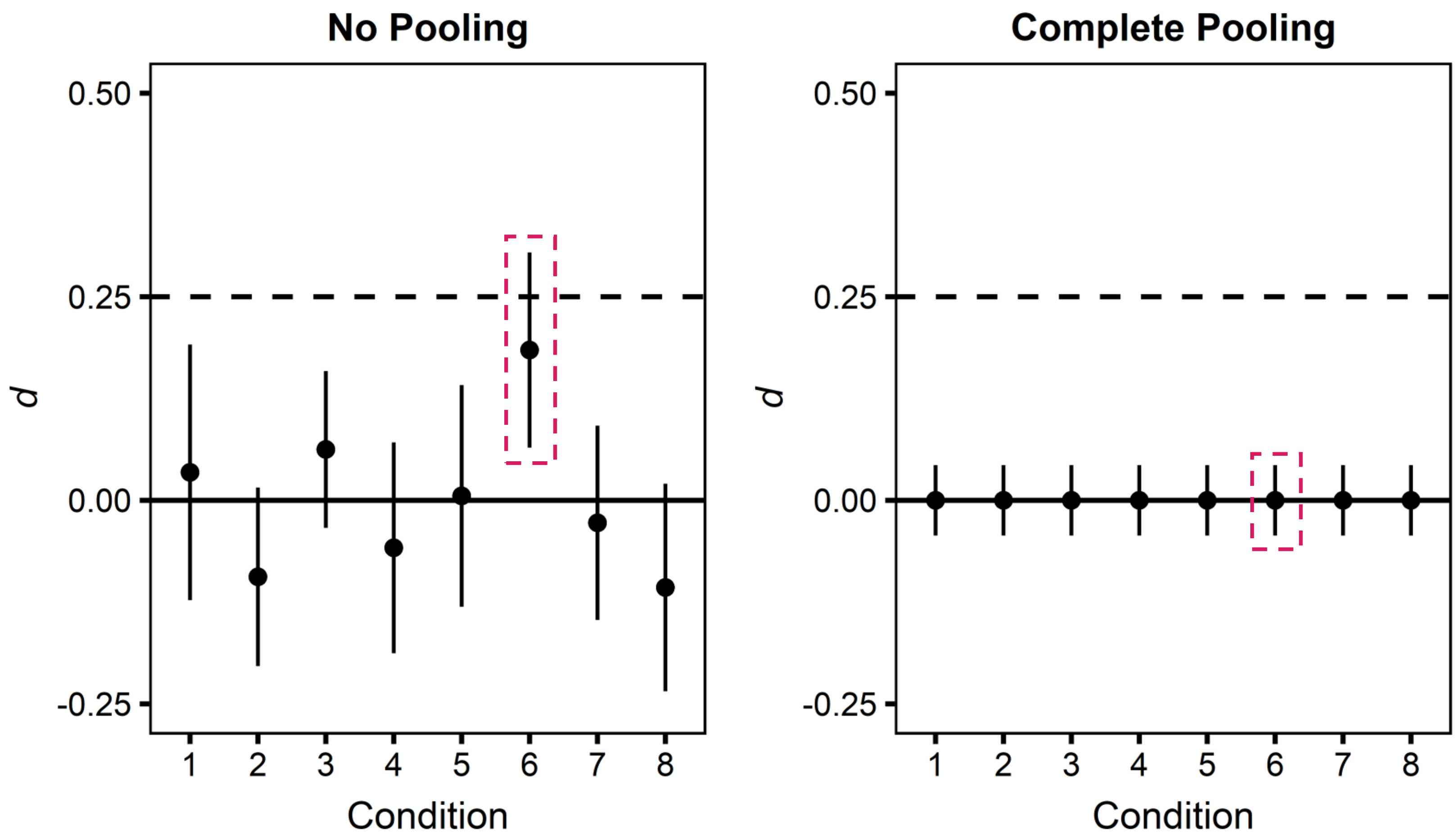
Why use multilevel models?

1. Data is nested.
   If we ignore nestedness, we misrepresent the data.

2. Ignoring nestedness does not work.
   If we pool across groups, we underestimate standard errors.

3. Using simple linear models does not work.


No Pooling
Complete Pooling
0.50
0.25
0.00
-0.25
d
1
2
3
4
5
6
7
8
Condition


No Pooling
Complete Pooling
0.50
0.25
0.00
-0.25
d
1 2 3 4 5 6 7 8
Condition

Why use multilevel models?

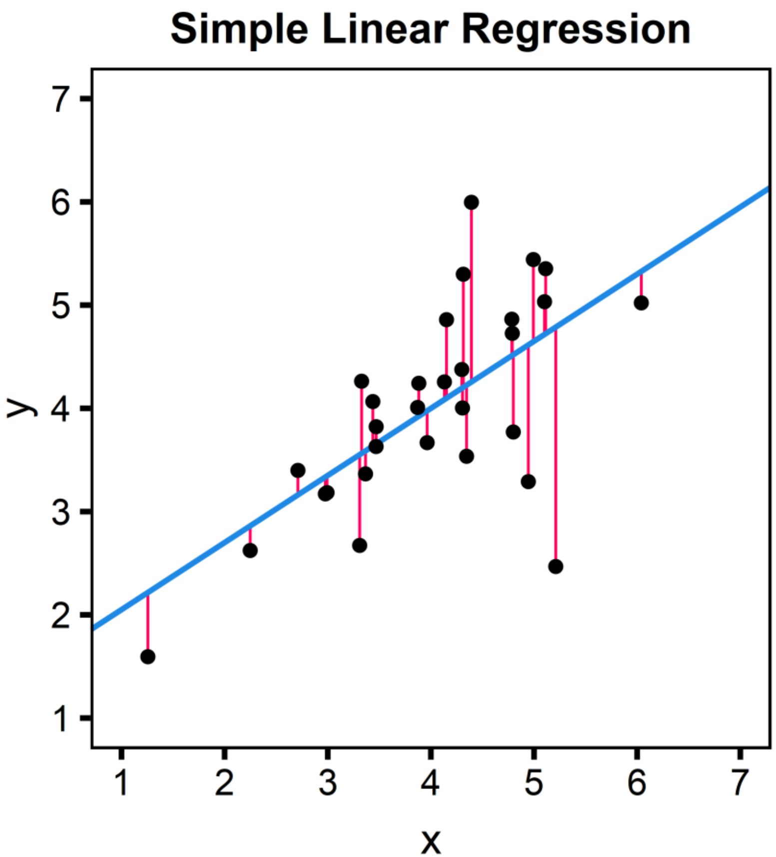
1. Data is nested.
   If we ignore nestedness, we misrepresent the data.

2. Ignoring nestedness does not work.
   If we pool across groups, we underestimate standard errors.

3. Using simple linear models does not work.
   If we do not pool across groups, we run into multiple-comparison problems.

What are multilevel models?

1. An extension of simple linear models.

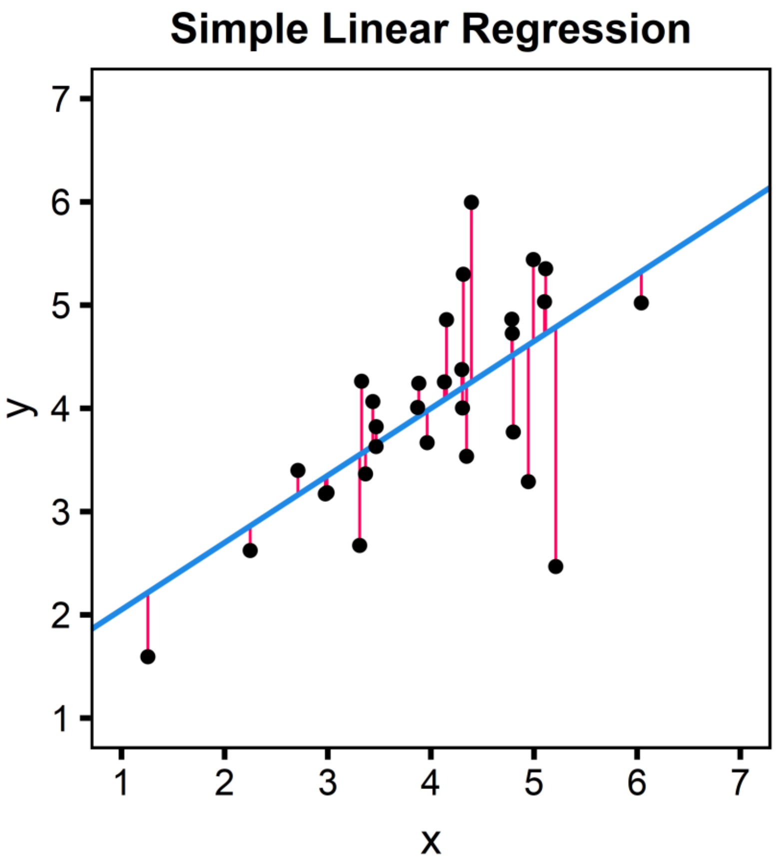


Intercept. Predicted value of y when $x = 0$.

Slope. Predicted change in y for each unit increase in y.

$$y_i \approx \boxed{\alpha} + \boxed{\beta} x_i$$

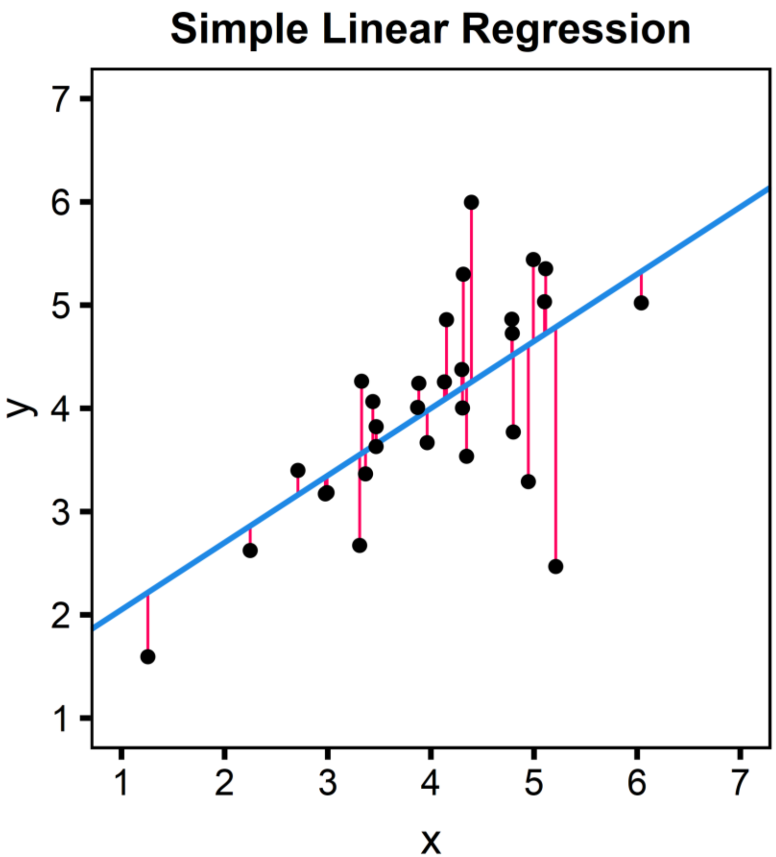
Intercept. Predicted value of y when $x = 0$.



Slope. Predicted change in y for each unit increase in y.

$$y_i = \boxed{\alpha} + \boxed{\beta} x_i + \boxed{e_i}$$

Residual Error. Difference between predicted and observed value of y_i for participant i.

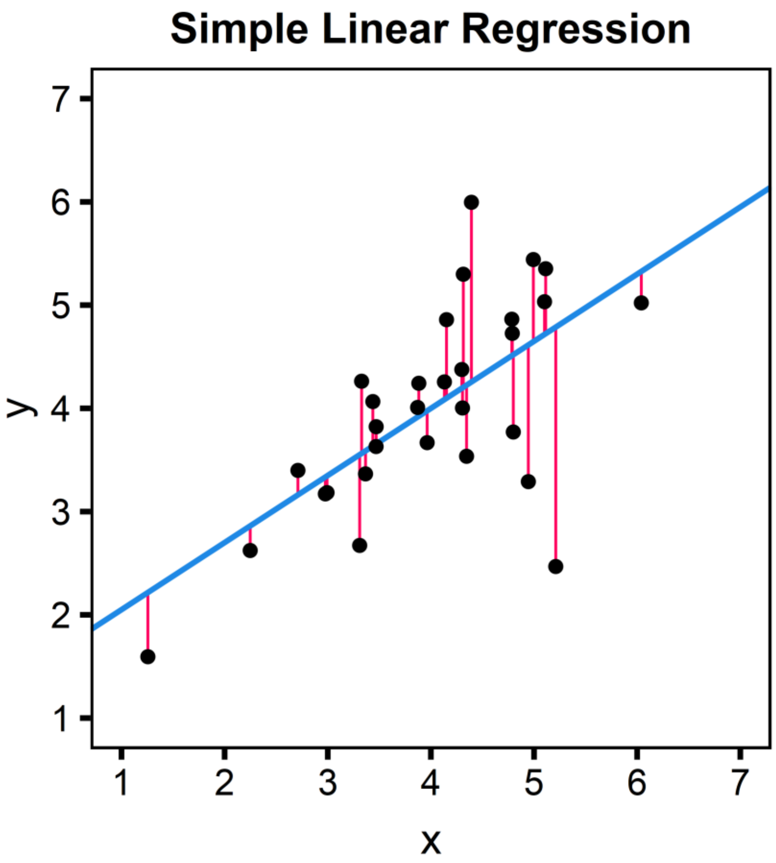


Residual Error. Difference between predicted and observed value of y_i for participant i.

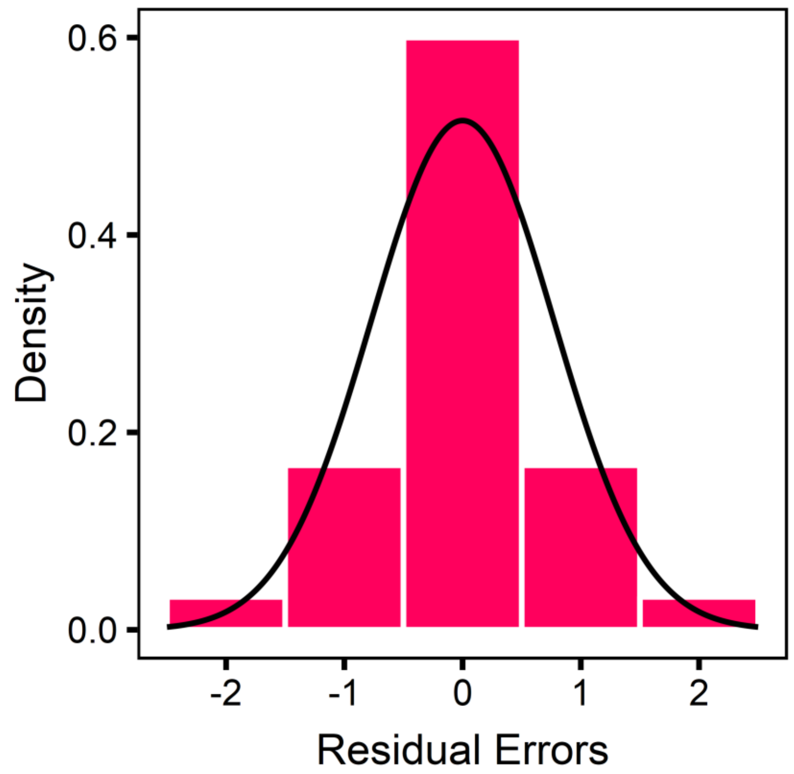
$$y_i = \alpha + \beta x_i + \boxed{e_i}$$

$$\boxed{e_i} \sim \mathrm{Normal}(0, \boxed{\sigma_I})$$

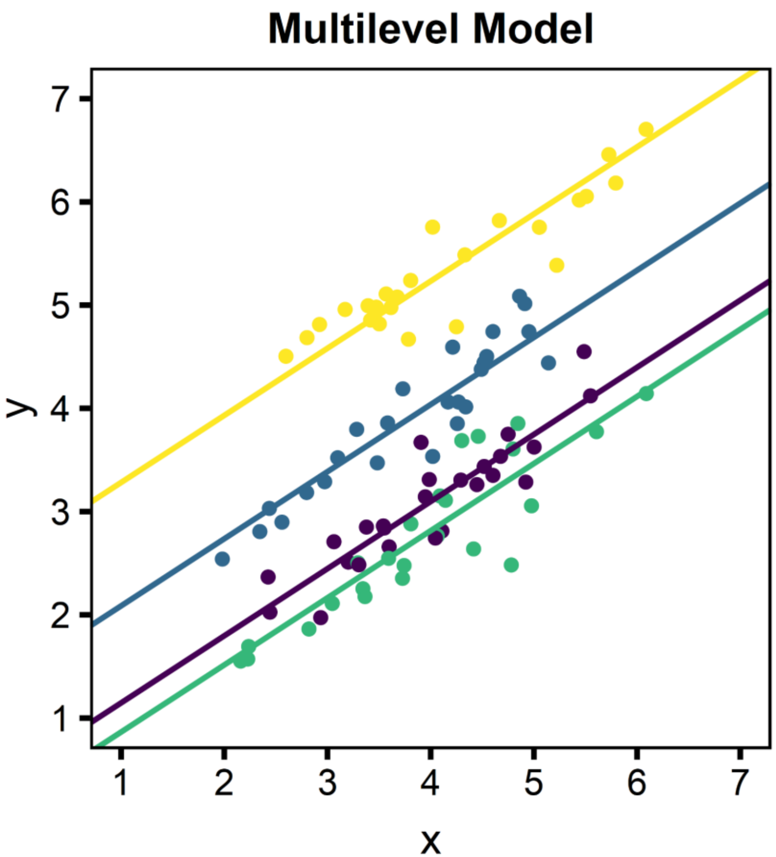
Residual Variance. e_i is centred around mean $\mu = 0$ and standard deviation σ_I.



$$e_i \sim \mathrm{Normal}(0, \sigma_I)$$



Varying (Random) Intercept.
Intercept for participants in group *j*.



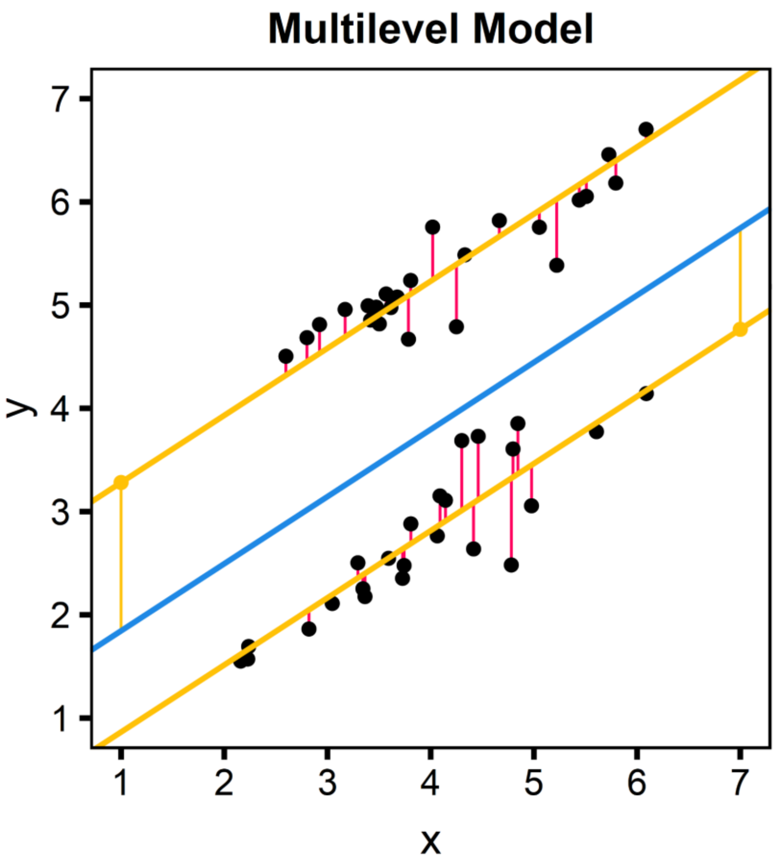
$$y_{ij} = \alpha + \boxed{\alpha_j} + \beta x_i + e_i$$

$$e_i \sim \mathrm{Normal}(0, \sigma_I)$$

$$\alpha_j \sim \mathrm{Normal}(0, \boxed{\sigma_J})$$

Between-group variance. α_j is centred around 0 with standard deviation σ_J.

Varying (Random) Intercept.
Intercept for participants in group *j*.



$$y_{ij} = \alpha + \boxed{\alpha_j} + \beta x_i + e_i$$

$$e_i \sim \mathrm{Normal}(0, \sigma_I)$$

$$\alpha_j \sim \mathrm{Normal}(0, \boxed{\sigma_J})$$

Between-group variance. α_j is centred around 0 with standard deviation σ_J.

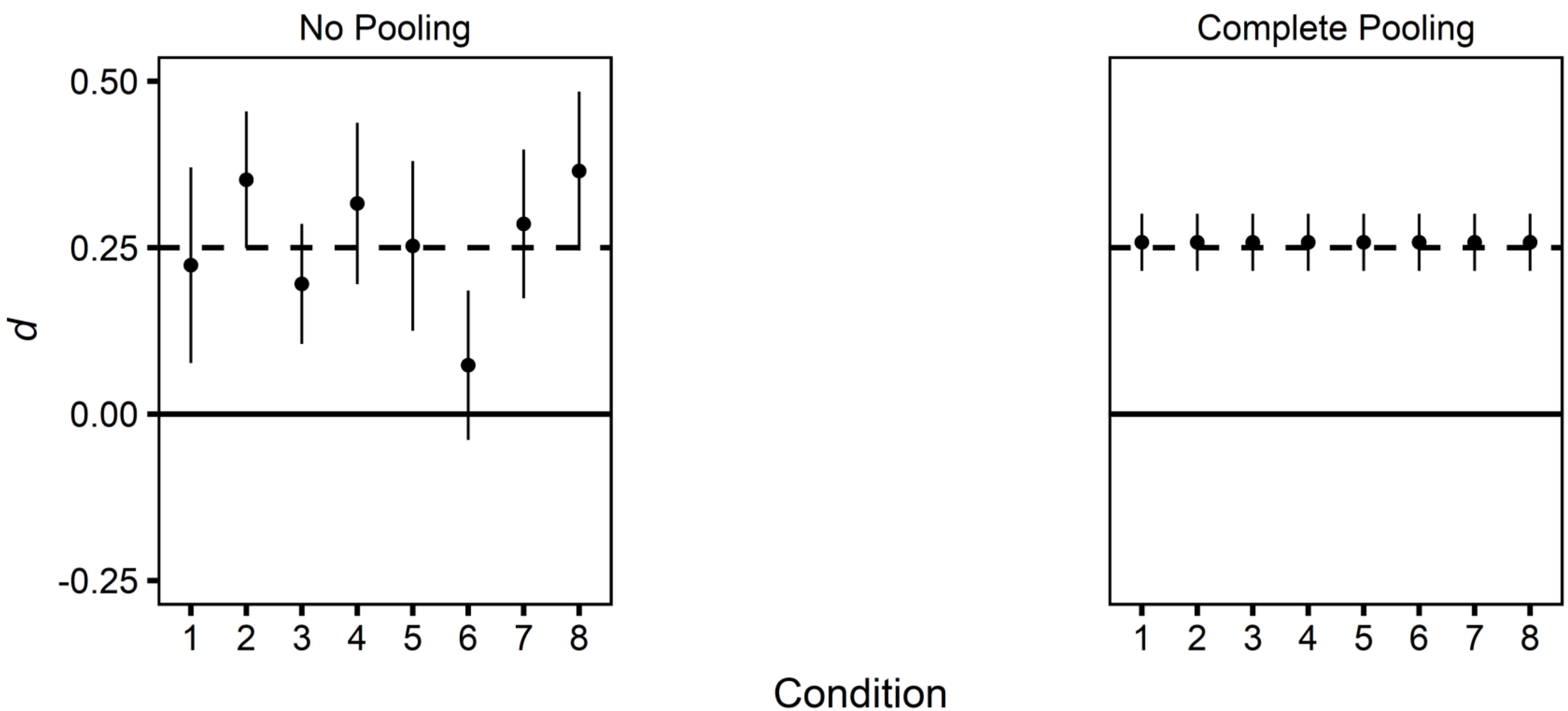
What are multilevel models?

1. An extension of simple linear models.
   Multilevel models add varying (random) intercepts and/or slopes to the regression equation.

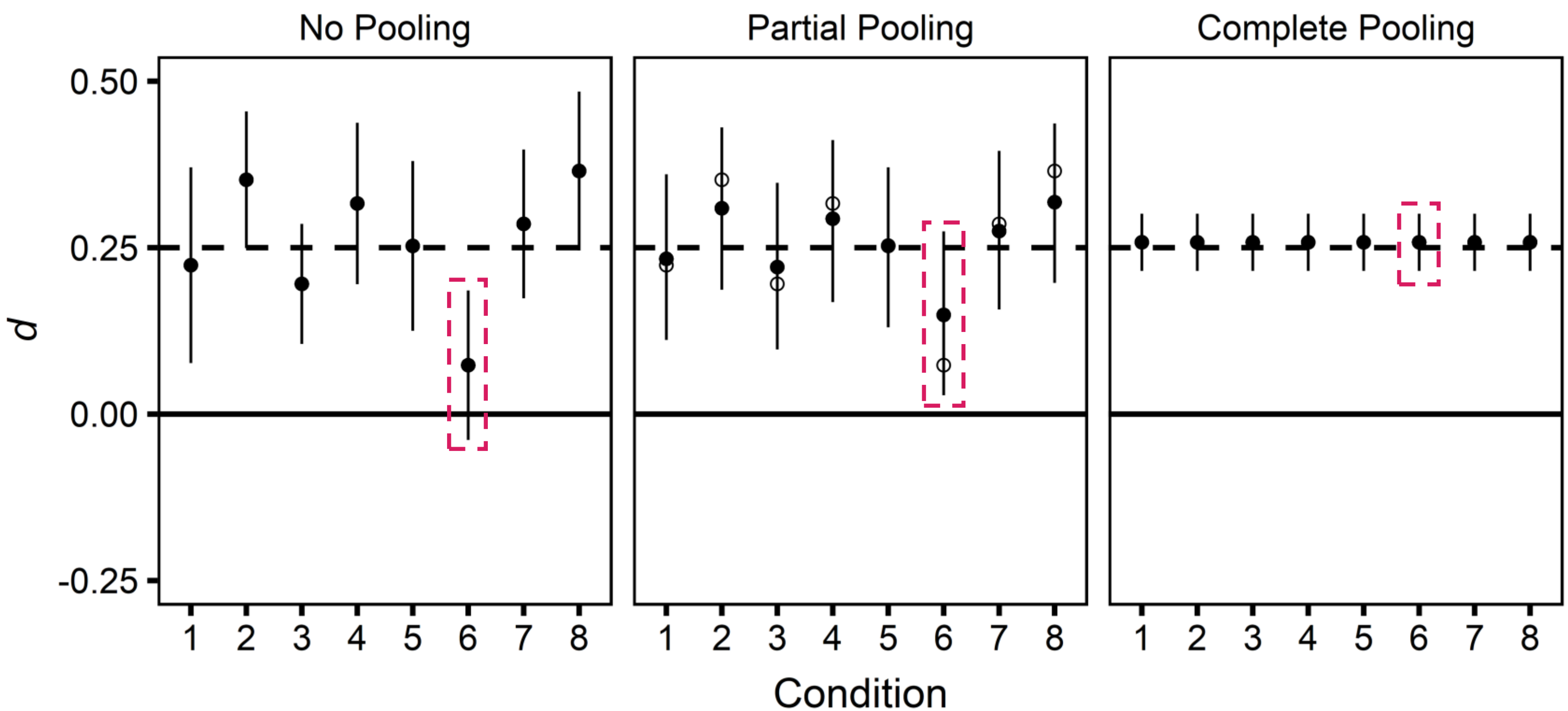
What are multilevel models?

1. An extension of simple linear models.
   Multilevel models add varying (random) intercepts and/or slopes to the regression equation.

2. A method for partial pooling.



Gelman, Hill, & Yajima, 2012


No Pooling
Partial Pooling
Complete Pooling
0.50
0.25
0.00
-0.25
d
1 2 3 4 5 6 7 8
Condition

What are multilevel models?

1. An extension of simple linear models.
   Multilevel models add varying (random) intercepts and/or slopes to the regression equation.

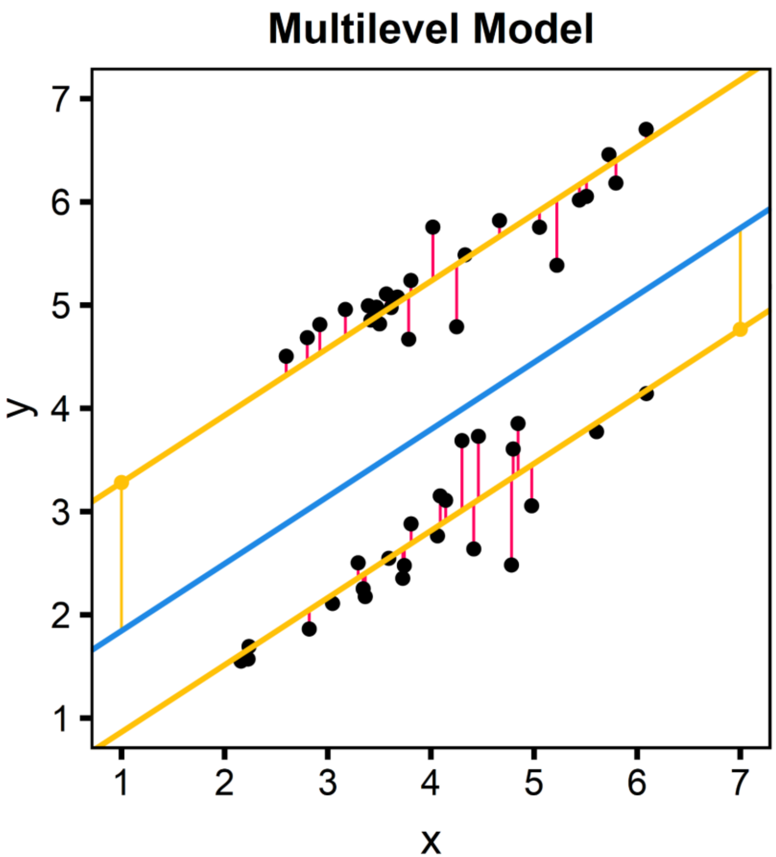
2. A method for partial pooling.
   Multilevel models shrink group-wise estimates toward each other.

What are multilevel models?

1. An extension of simple linear models.
   Multilevel models add varying (random) intercepts and/or slopes to the regression equation.

2. A method for partial pooling.
   Multilevel models shrink group-wise estimates toward each other.

3. A method to estimate variance.



$$e_i \sim \text{Normal}(0, \sigma_I)$$
$$\alpha_j \sim \text{Normal}(0, \sigma_J)$$

$$\text{ICC}_I = \frac{\sigma_I}{\sigma_I + \sigma_J} = 6\%$$

$$\text{ICC}_J = \frac{\sigma_J}{\sigma_I + \sigma_J} = 94\%$$

Intraclass Correlation Coefficient.
Proportion of total (unexplained) variance at each cluster level.

What are multilevel models?

1. An extension of simple linear models.
   Multilevel models add varying (random) intercepts and/or slopes to the regression equation.

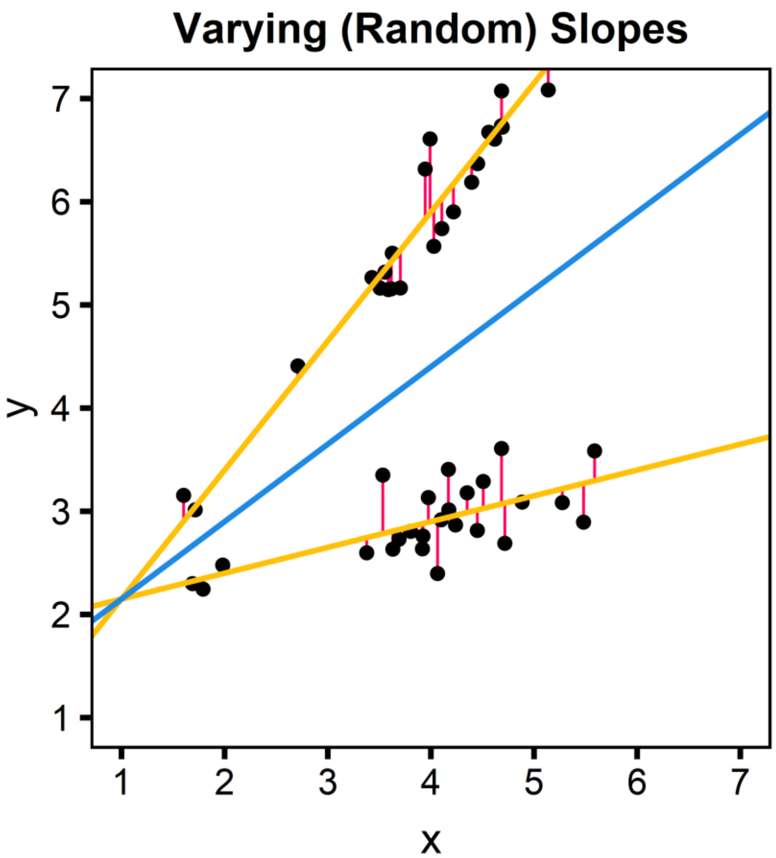
2. A method for partial pooling.
   Multilevel models shrink group-wise estimates toward each other.

3. A method to estimate variance.
   Multilevel models allow us to compare unexplained variance across levels.

How to estimate multilevel models?

```
> library(lme4)
> lmer(y ~ 1 + x + (1|jj), data = d)
```

$$y_{ij} = \boxed{\alpha} + \boxed{\alpha_j} + \boxed{\beta} x_i + e_i$$



Varying (Random) Slope.
Slope for participants in group *j*.

$$y_{ij} = (\alpha + \alpha_j) + (\beta + \boxed{\beta_j})x_i + e_i$$

How to estimate multilevel models?

```
> library(lme4)
> lmer(y ~ 1 + x + (1 + x|jj), data = d)
```

$$y_{ij} = (\boxed{\alpha} + \boxed{\alpha_j}) + (\boxed{\beta} + \boxed{\beta_j})x_i + e_i$$

Workshop

```
> d <- read_rds("data/d_wk6.rds")
```

`ii`	Each participant (*N* = 900) has a unique identification number (1 ≤ `ii` ≤ 900).
`jj`	Each neighbourhood (*J* = 30) has a unique identification number (1 ≤ `jj` ≤ 30).
`contact`	Each participant reported the amount of contact with immigrants they had had (1 = *None*, 7 = *A lot*).
`diversity`	Each neighbourhood had a known proportion of (first-, second-, and third-generation) immigrants.
`attitudes`	Each participants reported how they felt about immigrants in general (0 = Cold, 100 = Warm).

Consequences of Diversity for Social Cohesion and Prejudice: The Missing Dimension of Intergroup Contact

Miles Hewstone
University of Oxford

A controversial claim that diversity has negative consequences for trust and other outcomes spawned a contentious debate in sociology and political science, but was hardly noted in social psychology. I summarize the debate, and argue that the efficacy of direct, face-to-face intergroup contact as a means of reducing prejudice is a stark omission, as I illustrate with evidence of the association between diversity, on the one hand, and trust, prejudice, and social capital on the other. I also consider two other contributions of contact theory to this issue, namely that contact with members of one group has an impact on attitudes toward members of other groups; and that contact should be studied via social networks. Despite the importance I attach to contact, I note two "enemies of contact," resegregation in ostensibly desegregated settings, and negatively valenced contact. Finally, I point to the kind of research we should do, in order to increase the impact of our work on the public policy debate on this issue.

Hewstone, 2015, https://doi.org/10.1111/josi.12120

When should we use multilevel models?

1. When data is nested.

   If we know of any clustering in the way the data was collected or generated, we should use them.

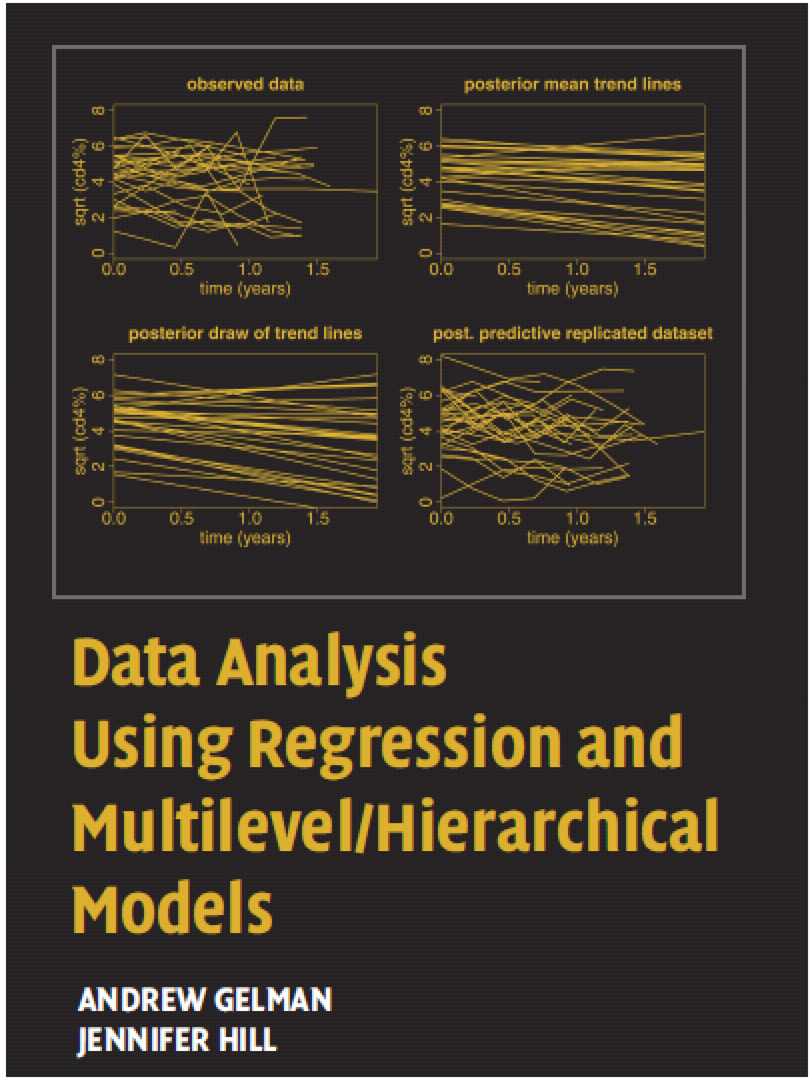
2. When there is little variance between groups.

   When ICCs are low, we should still use them. Multilevel models include simple regression as a special case.

3. When we have few participants/groups.

   As long as we have more than two observations in some or most clusters, we should use multilevel models (Gelman & Hill, 2007). Power, of course, will suffer.

http://elevanth.org/blog/2017/08/24/multilevel-regression-as-default/


observed data
posterior mean trend lines
posterior draw of trend lines
post. predictive replicated dataset
sqrt (cd4%)
time (years)
Data Analysis Using Regression and Multilevel/Hierarchical Models
ANDREW GELMAN
JENNIFER HILL