Experimental Methods: Clustering and Spillovers
Constantine Manda | Yale University

2016 Impact evaluation workshop & EAST AFRICA Evidence Summit
Which level of randomization?

**INDIVIDUAL** unit:
- e.g. pupil, mother, father, household

**CLUSTER** unit:
- e.g. class, health clinic, village
Which level of randomization?

Individual or Cluster?
Which level of randomization?

Individual
Which level of randomization?

Cluster (group)
Which level of randomization?

For example, school
Which level of randomization?
Factors that matter

1. Unit of Implementation
2. Spillovers
3. Sampling unit and statistical power
4. Ethics and Fairness
Which level of randomization?

Unit of randomization should be at least at the level of the unit of program intervention:

Example:

IF program targets *classes* ->

Randomization may be at class, school, community level
Which Level of Randomization?

A. Unit of Implementation

B. Spillovers: individual in the control also benefits from the treatment

C. Sampling and Cost

D. Ethics and Fairness
With spillovers over/under-estimate impact.
Higher unit of randomization reduces risk of spillovers

Example: randomize at school level to reduce interaction between treatment and control pupils
Question: Is there risk of spillovers in the following programs?

• Example 1: Job training program

• Job seekers in the control are less likely to get jobs as a result of the program, if the available jobs are given to those who received training.

• Measured program impact = jobs in treatment – jobs in comparison

• As a result, measured program impact is higher than the real program impact
Question: Is there risk of spillovers in the following programs?

- Example 2: **HIV-AIDS education program**
- People who attend education program will tell their friends
- Measured program impact = HIV-aids awareness in treatment – HIV-aids awareness in comparison
- As a result, *measured* impact is lower than the *real impact*
Which Level of Randomization?

A. Unit of Implementation
B. Spillovers
C. Sampling unit and statistical power
D. Ethics and Fairness
Higher unit of randomization requires larger sample

Same statistical power:

School randomization:
- 800 in treatment and 800 in control in
- 40 treatment and 40 control schools, 20 pupils per school

Pupil randomization:
- 393 in treatment and 393 in control

Assume intra-cluster correlation of 0.05
Need to increase number of clusters, not number of individuals per cluster

For example, the following two studies have the same power:

• 80 clusters, 20 individuals per cluster
• 40 clusters, 1067 individuals per cluster

That’s 1,600 individuals compared to 42,680!

Assume intra-cluster correlation of 0.05
At what level should we randomize for a school deworming program?

A. School  
B. Teacher  
C. Classroom  
D. Clinic  
E. Community  
F. Pupil
At what level should we randomize for a conditional cash transfer program?

A. Individual  
B. Household  
C. Village/Community
At what level should we randomize for a teacher training program?

A. School
B. Teacher
C. Classroom
D. Pupil
At what level should we randomize for a small-business start-up grant?

A. Individual
B. Household
C. Clinic
D. School
E. Village/Community