Intro to Mediation

Session 20

Ben Kelcey
University of Cincinnati

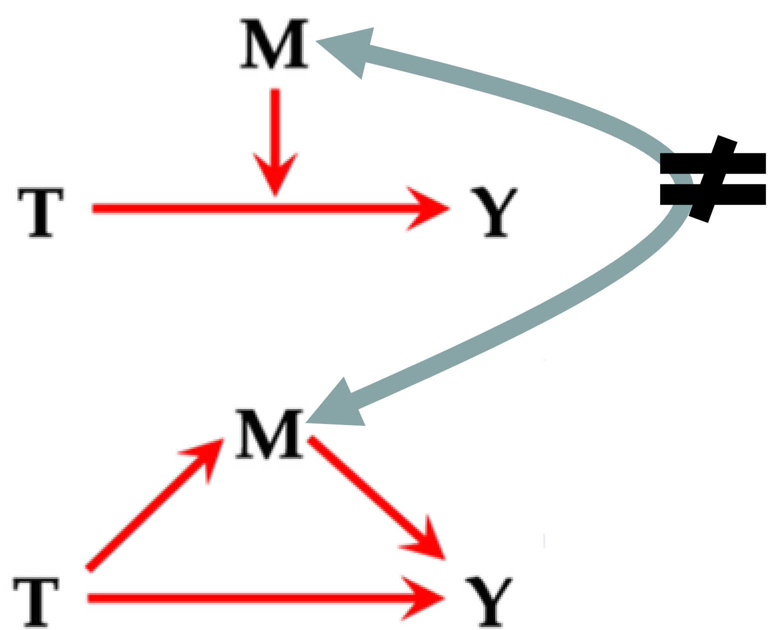
Sessions 20 & 21

- Introduction to mediation
- Analysis of mediation studies
- Design of mediation studies

Main, Moderation, Mediation

- Main

- Moderation
- Mediation

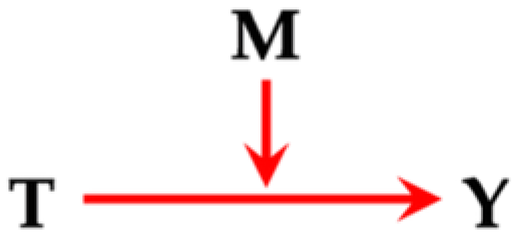


Main Effects

$$T \longrightarrow Y$$

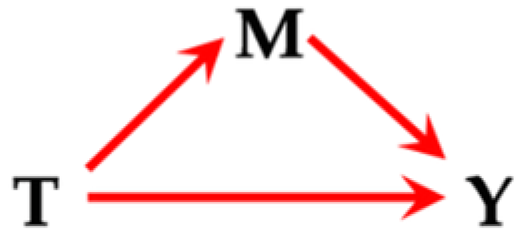
- Most studies initially focus on main or average effects
  - Main effects describe whether a program works on average
- Main effects studies are limited in their capacity to comprehensive assessments and evidentiary bases from which to draw inferences

Moderation



- Do effects vary by subgroup or context?
- A moderating variable is a pretreatment variable that interacts with the treatment such that the impact of the treatment depends on the value of the moderator variable
- Moderation addresses for whom and under what circumstances a treatment is effective

Mediation



- Mediation analyses unpack the pathways or intermediate variables through which the treatment operates on the outcome
  - Tests a theory of action
  - Informs how a treatment works and identifies any breaks in a theory

Moderation v. Mediation

- Moderation: heterogeneity in treatment effects across subgroups and conditions/contexts
  - Introduces interaction between treatment and pretreatment variables to probe differential effects
- Mediation: probes the mechanisms through which a treatment impacts an outcome
  - Introduces intermediate variables (post-treatment but pre-outcome) that lie on the causal pathway from the treatment to outcome to probe the theory of action

Example Context

A seismic shift in national research priorities over the past 6 years has led to a dramatic increase in the number of large-scale randomized experiments designed to test the impact of educational interventions on student outcomes. Spybrook (2007) identified 55 such trials supported by the Institute for Education Sciences. Of these, the vast majority assigned groups, typically schools or classrooms rather than individuals, to interventions. The majority of the innovative interventions attempted to improve student learning by improving classroom teaching.

Raudenbush, S., & Sadoff, S. (2008). Statistical inference when classroom quality is measured with error. Journal of Research on Educational Effectiveness, 1(2), 138–154.

Context

A major aim of these studies is to evaluate the impact on student learning of assignment to an innovative classroom intervention. This aim can be achieved, in principle, without measuring the quality of classroom instruction. However, the interpretation of findings from such a study will typically be ambiguous.

Context

Consider a study in which the assignment of schools or classrooms to a novel instructional innovation is found to have no significant impact on student learning. Assume that the study design was unbiased and provided adequate statistical power to detect a nonnegligible effect. Two explanations immediately arise. Program evaluators refer to these as "theory failure" versus "implementation failure" (Rossi, Lipsey, & Freeman, 2004).

Context

First, it may be that the innovation changed classroom instruction in the ways intended but that those classroom changes made no difference in student learning. The term *theory failure* describes this scenario because the theory that links intended changes in instruction to intended student outcomes will have proven incorrect.

Second, the innovation may never have been effectively implemented in classrooms. Perhaps the innovators lacked skill in working with teachers or perhaps the teachers lacked the skill, knowledge, or motivation to put the innovative ideas to work in their teaching. In any case, program theory about the relationship between the intended instruction and student outcomes was never tested, leading to "implementation failure."

UNIVERSITY OF Cincinnati

Context

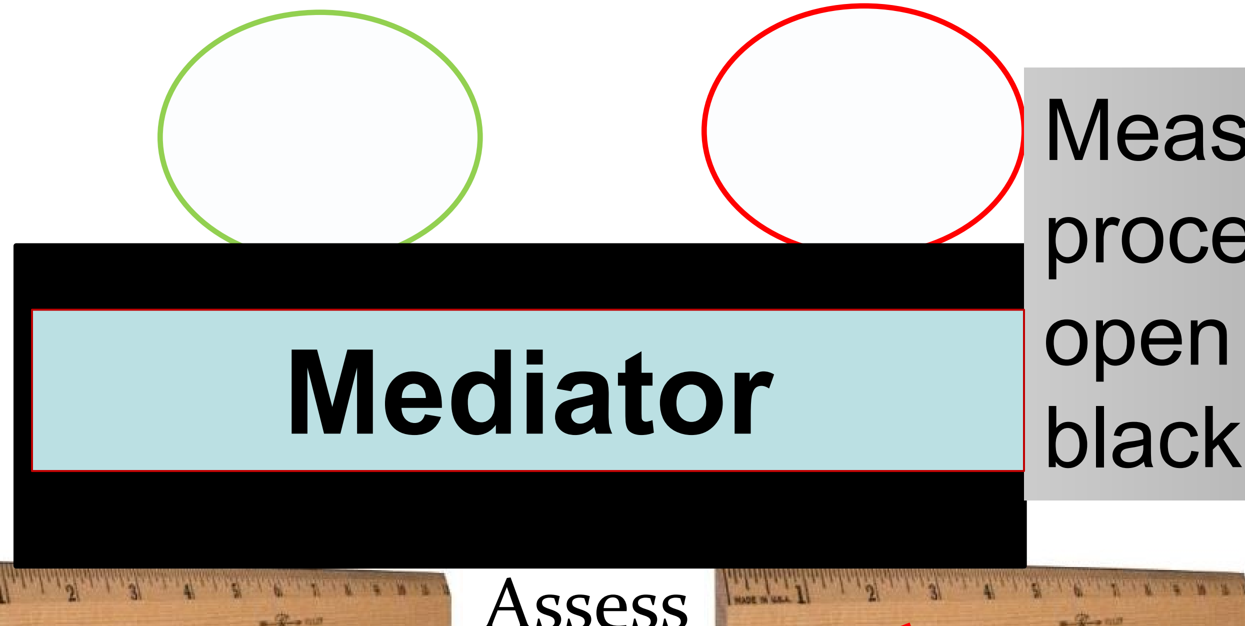
Without valid assessments of instructional process, it would be impossible to distinguish between these two explanations, severely limiting the study's contribution to knowledge. One would never know whether the theory underlying the program had in fact been tested.

Suppose instead that assignment to the innovation did produce gains in student learning. One might then assume that the innovation "worked" by

improving instruction in the ways the program designers intended. But without valid measurements of instruction, this conclusion would be unwarranted. Perhaps the innovation "worked" in other ways, an assertion that could not be probed without studying the impact of the innovation on instruction. Once again, a failure to measure key aspects of classroom life yields major ambiguities in the findings.

Mediation

Sample of Schools



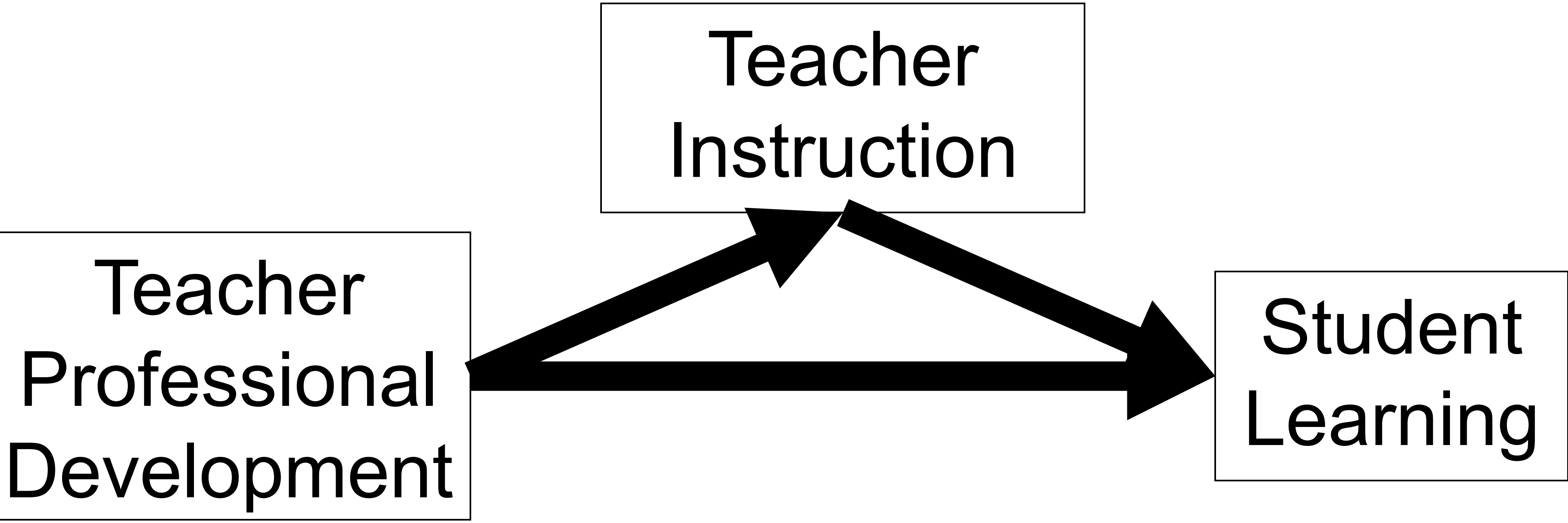
Measuring the process helps to open up the blackbox

Compare

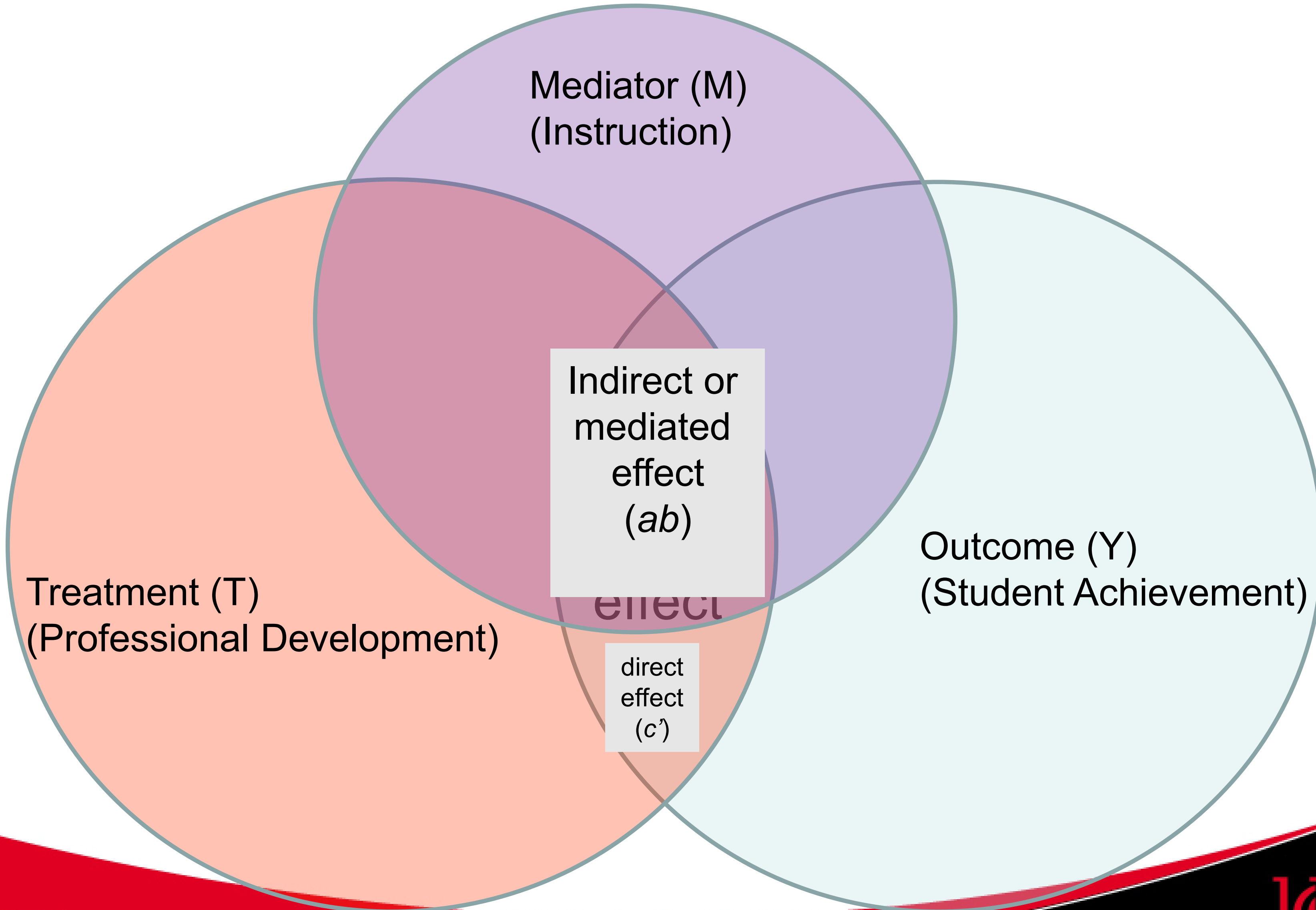
Inf
int

Theory failure or implementation failure?

Example


Teacher Instruction
Teacher Professional Development
Student Learning

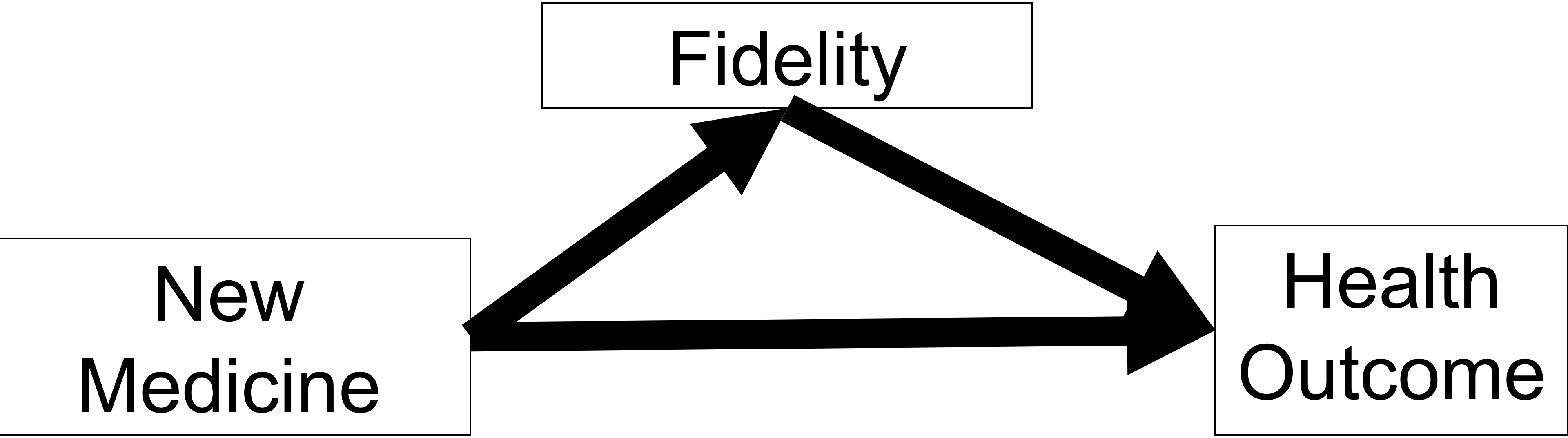
Mediation


Mediator (M)
(Instruction)
Indirect or mediated effect (ab)
Treatment (T)
(Professional Development)
Outcome (Y)
(Student Achievement)
direct effect (c')

Examples

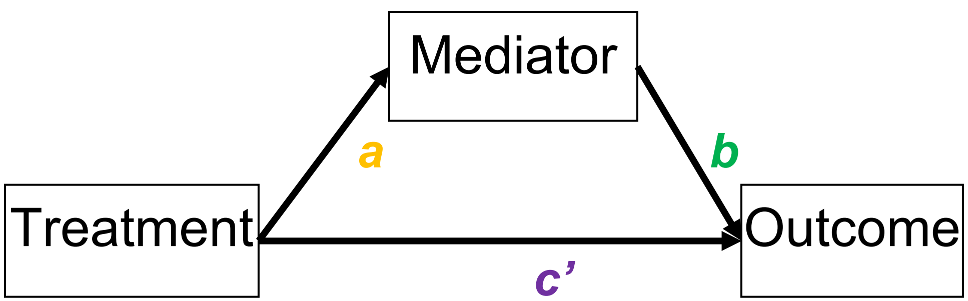
- Whole school intervention → *Instruction* → Achievement
- Therapy → *Engagement* → Outcome
  - E.g., Patient engagement
- Treatment → *Attitude* → Outcome
  - E.g., Motivational interviewing
- Treatment → *Fidelity* → Outcome
  - E.g., Patient fidelity to treatment with side effects
- Professional Development → *Knowledge* → *Instruction* → Achievement

Single Level Example



Simple single level mediation

$$Y_i = \beta_0 + c T_i + e_i$$



$$M_i = \beta_0 + a T_i + e_i$$

$$Y_i = \beta_0' + b M_i + c' T_i + e_i'$$

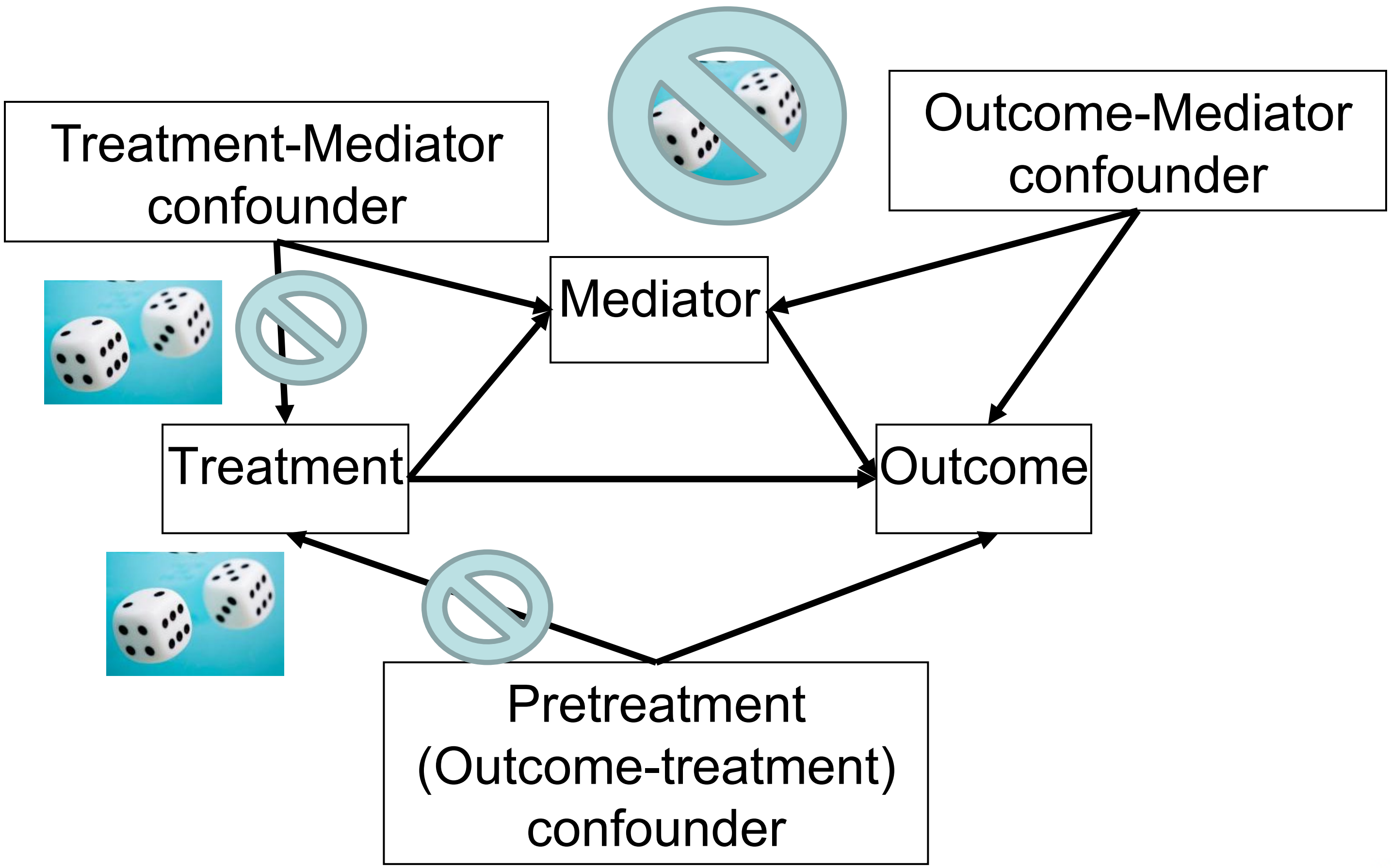
total effect = indirect effect + direct effect

c = ab + c'

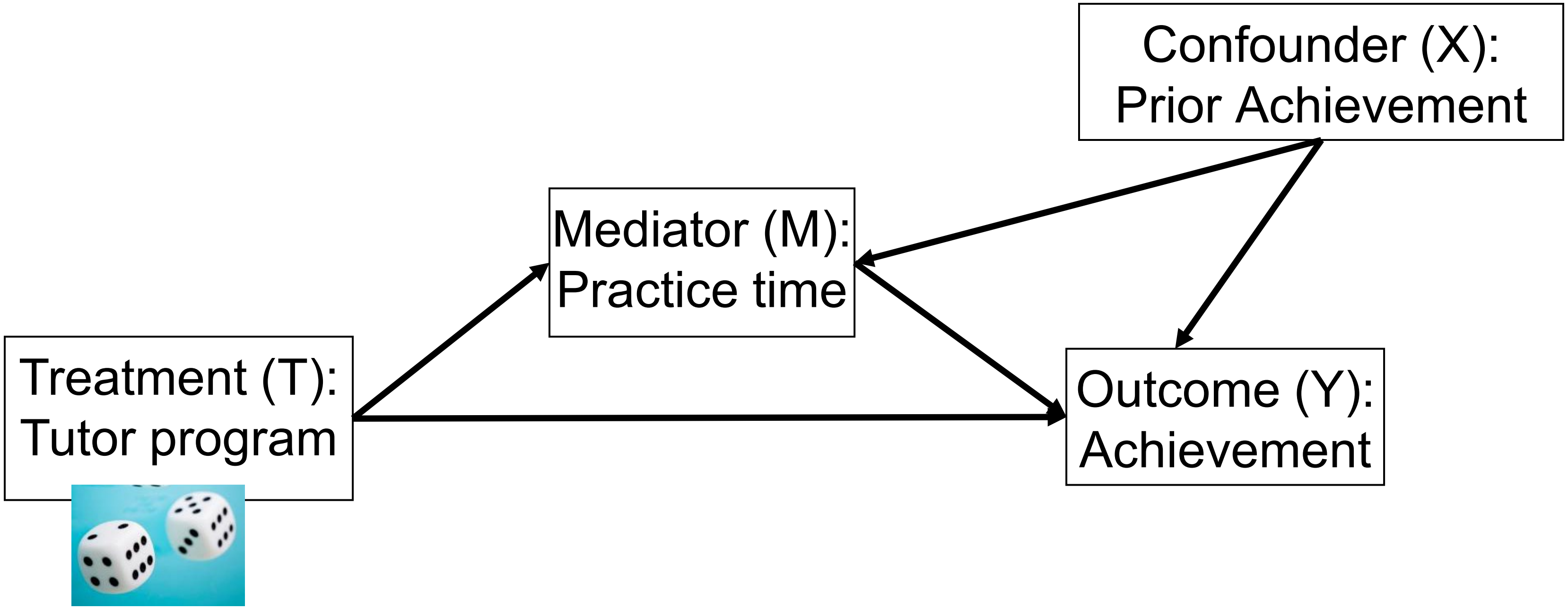
Assumptions: Sequential Ignorability

- Historical literature has generally treated estimates of mediation as causal
  - Causal inference regarding mediation requires that BOTH the treatment and mediator be randomly assigned or are ignorable
  - Random assignment of treatment ensures there are no pretreatment confounders that explain the observed outcome differences
  - Random assignment of the mediator ensures there no outcome-mediator confounders

Sequential Ignorability



Sequential Ignorability



$$M_i = \beta_0 + aT_i \qquad\qquad + e_i$$

$$Y_i = \beta_{0Y} + bM_i + c'\ T_i \qquad\qquad + e_{iY}$$

Assumptions: Correct Specification

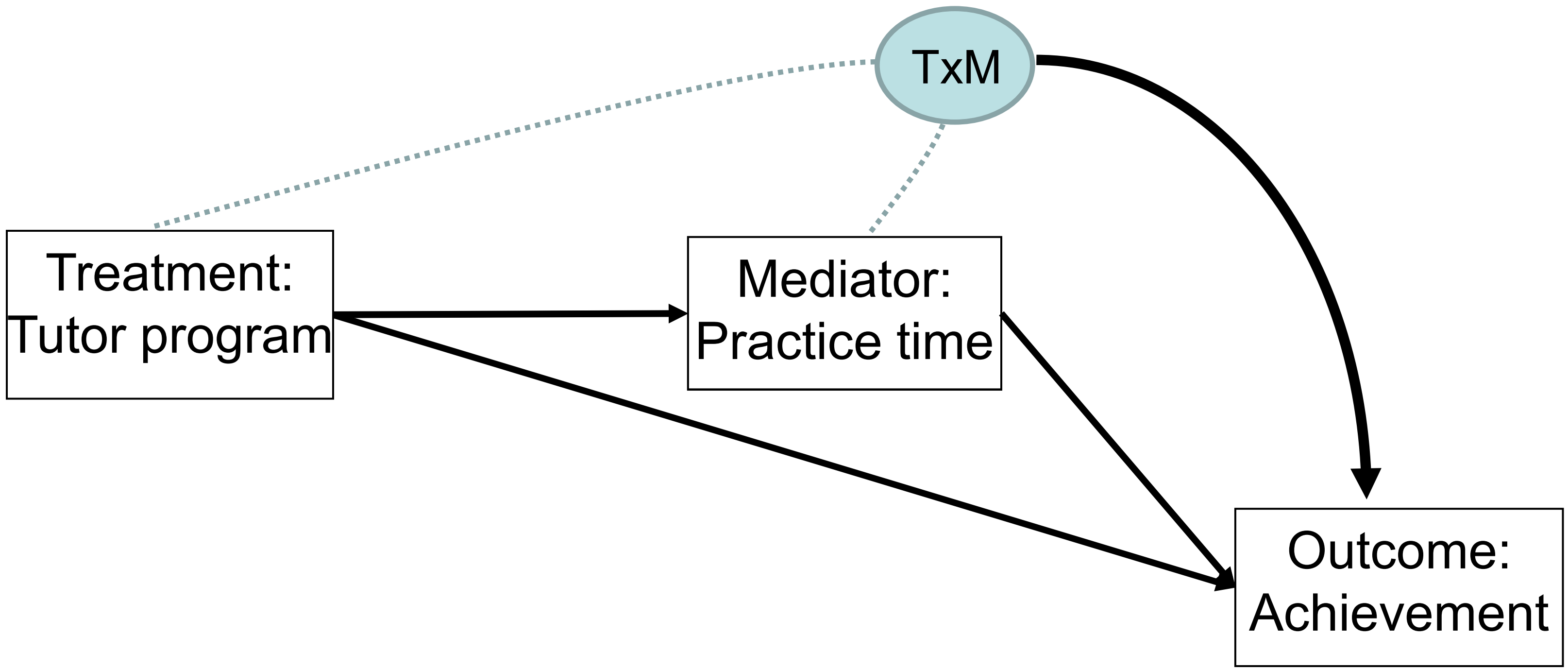
- Our models of mediation often differ from the truth
- Just because we estimate the model

$$T \to M \to Y$$

does not ensure that the relationships are causal

  - Inferences are sensitive and susceptible to specification bias

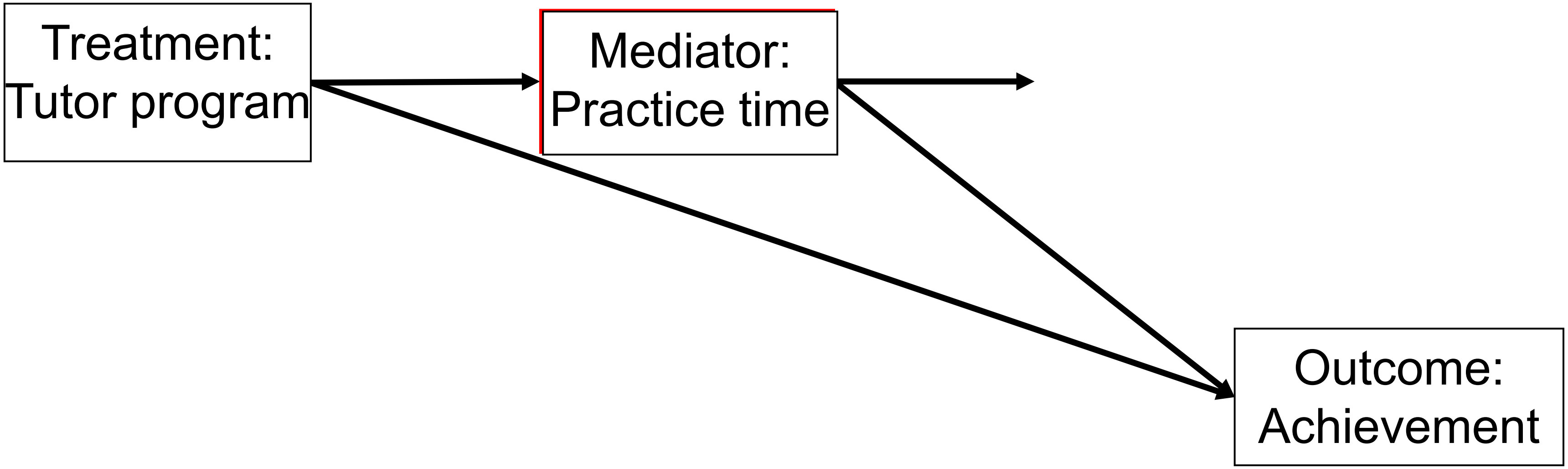
Correct Specification



$$M_i = \beta_0 + aT_i + e_i$$

$$Y_i = \beta_{0Y} + bM_i + c' \; T_i + \beta_1 M_i T_i + e_{iY}$$

Correct Specification

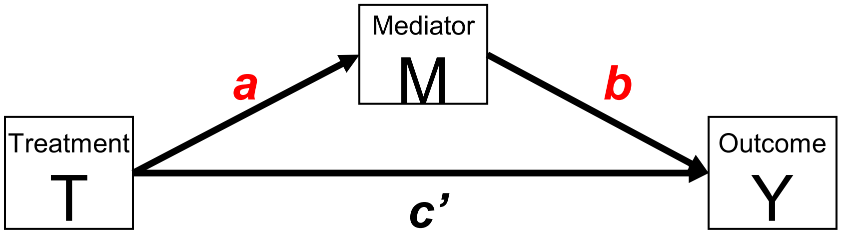


$$M_i^{(1)} = \beta_{01} + a_1 T_i + e_{i1}$$

$$M_i^{(2)} = \beta_{02} + b_1 M_i^{(1)} + a_2 T_i + e_{i2}$$

$$Y_i = \beta_{0Y} + c_2 M_i^{(2)} + c_2 M_i^{(1)} + c' \; T_i + e_{iY}$$

Testing for mediation effects



- ***ab*** is an estimate of the mediation effect (when there is not an interaction between T, M)
- Several possible approaches to test whether ab is nonzero and they differ in their power and type 1 error rates

Tests of Mediation

- Some Common Tests of mediation
  - Sobel test
  - Test of joint significance
  - Monte Carlo interval test
  - Bootstrap resampling

Sobel test

- Compare $z = \dfrac{ab}{se_{ab}}$ to a standard normal distribution
- Confidence interval around *ab*

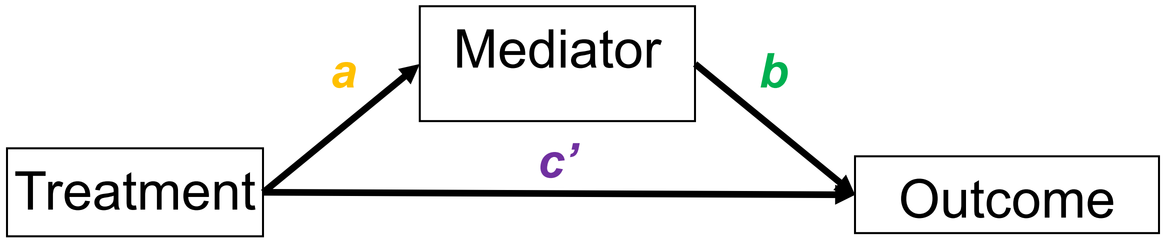
  CI = $ab \pm (z_{critical})(se_{ab})$

- Tends to be most conservative (lowest power and low type1 error rate relative to other tests) because distribution of ab is only approximately normal in large samples

Joint significance approach

1) Test a path
2) Test b path

*If both are significant then infer mediation



--Power is good approximation to more complex methods

--Type I error rate slightly lower than expected

Monte Carlo Interval Test

- Estimates empirical distribution of the ab product using resampling based methods (similar to parametric bootstrapping)
  - Draw samples of a and b from their respective distributions, multiply them and repeat to approximate the posterior distribution of ab
  - If CI does not include zero, then infer mediation
- Does not require full data (useful for design purposes)
- Does not assume the sampling distribution of the indirect effect is symmetric
- Typically found to be very powerful and comparable to bootstrap based intervals

Bootstrap resampling methods

- Estimates empirical distribution of the ab product using resampling based methods (similar to MC interval test)
- Several variations
  - Parametric percentile (resample residuals)
  - Non-parametric percentile (resample cases)
  - Bias-corrected versions of both (correct for difference between point estimate and median of empirical distribution)
- Typically found to be most powerful and accurate

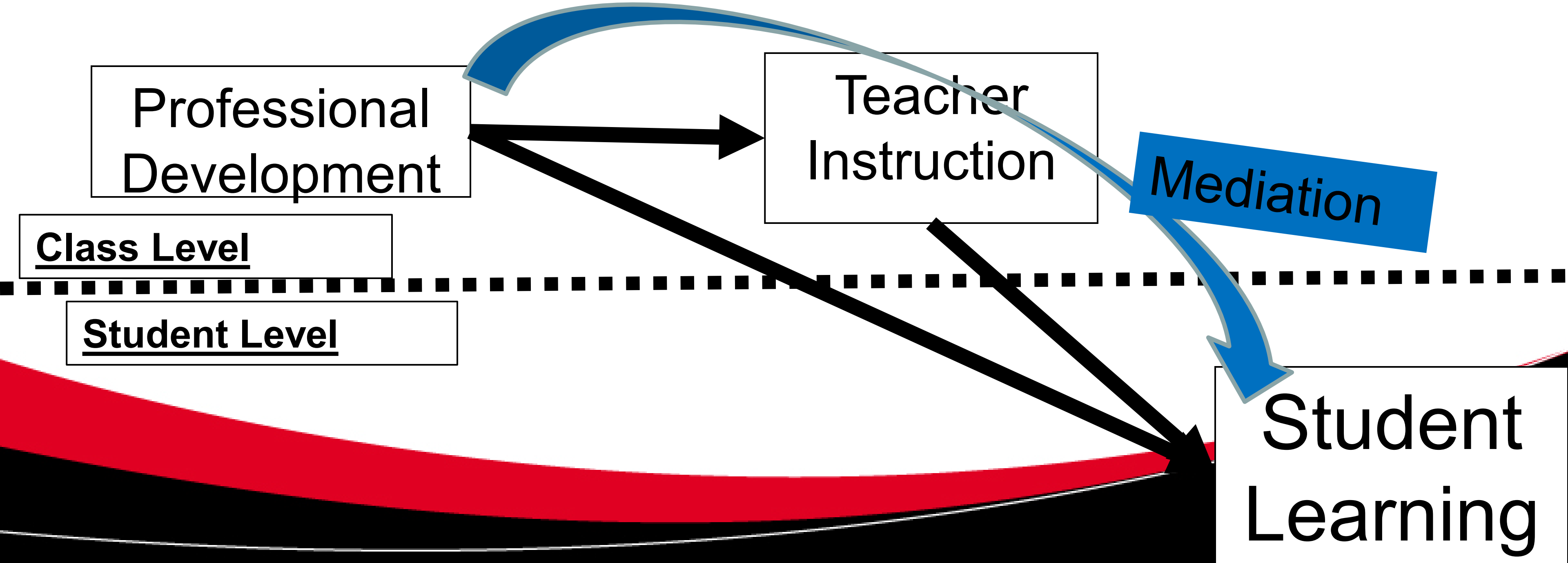
Mediation in
Cluster Randomized Trials

Mediation in Cluster Randomized Trials

- Lots of combinations and models depend on level at which each variable is assessed: $T \rightarrow M \rightarrow Y$, *e.g.,*
  - Upper-level mediation [$2 \rightarrow 2 \rightarrow 1$]
  - Cross-level mediation [$2 \rightarrow 1 \rightarrow 1$]
  - Three level and sequential mediation versions e.g., [$3 \rightarrow 2 \rightarrow 2 \rightarrow 1$] when considering how school randomized professional development programs impact student achievement via instruction (via teacher knowledge)

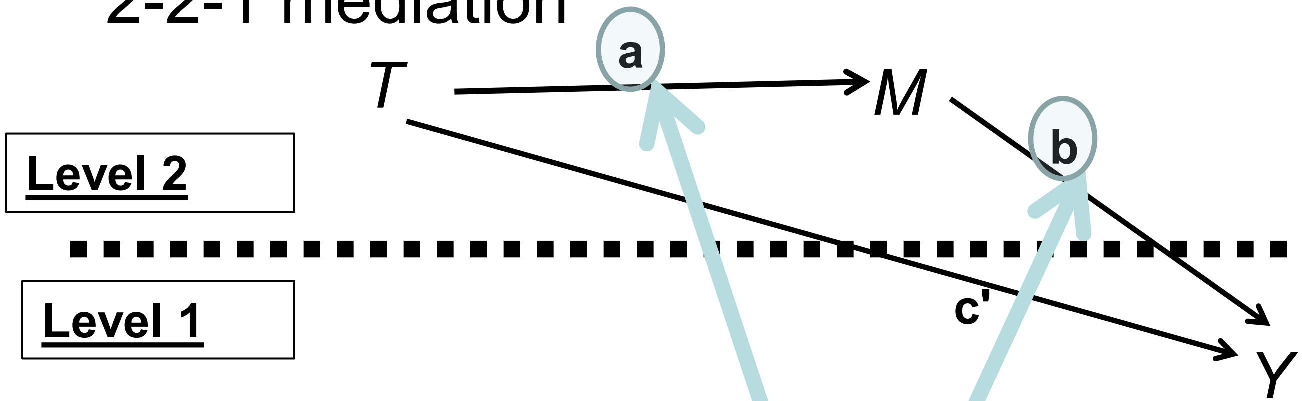
Simple 2-2-1 Mediation

- Imagine a cluster randomized trial that assigns teachers to receive different amounts of professional development (PD). The aim of this PD is to improve the teachers' quality of instruction (IQ) so that, in turn, students' achievement (Y) increases.



2-2-1 Mediation Model

2-2-1 mediation



Mediation model

Class level $$M_j = \pi_0 + aT_j + \varepsilon_j^M \qquad \varepsilon_j^M \sim N(0, \sigma_{M|}^2)$$

Outcome model

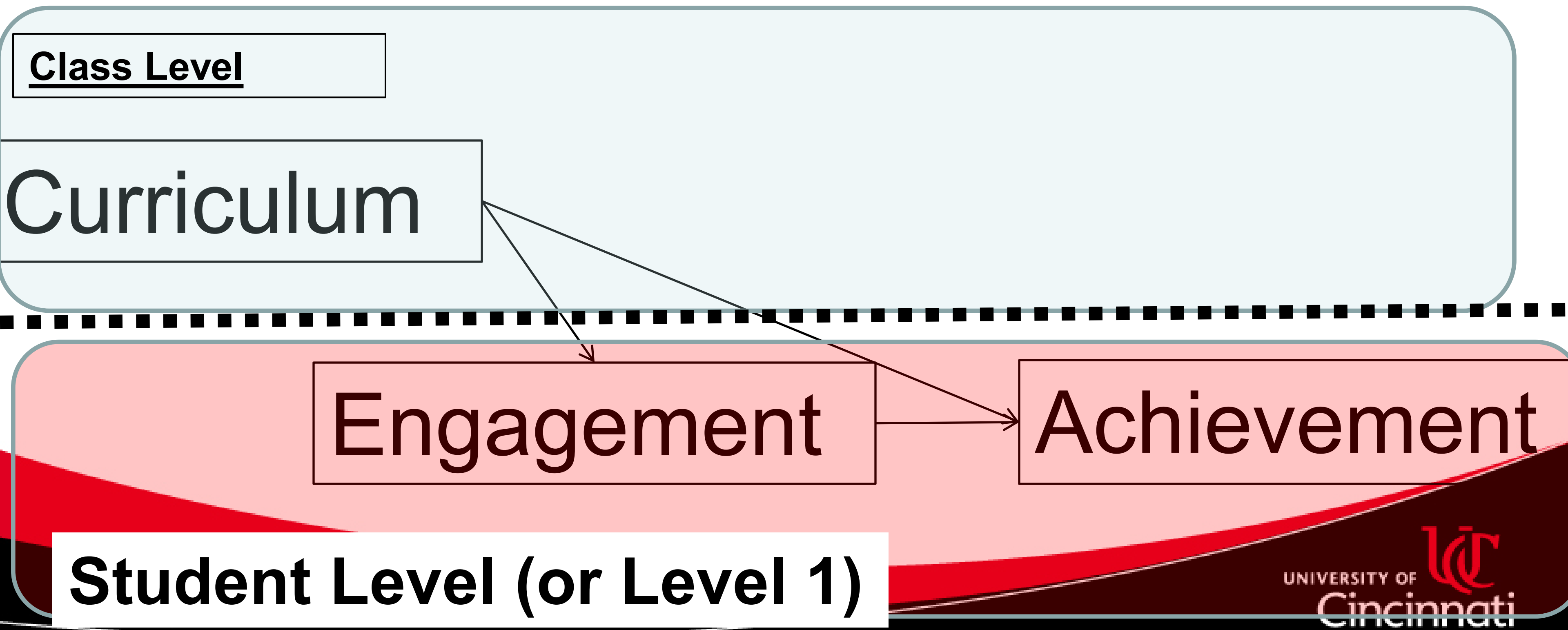
Student level $$Y_{ij} = \beta_{0j} + \varepsilon_{ij}^Y \qquad \varepsilon_{ij}^Y \sim N(0, \sigma_{Y|}^2)$$

Class level $$\beta_{0j} = \gamma_{00} + bM_j + c'T_j + u_{0j} \qquad u_{0j} \sim N(0, \tau_{Y|}^2)$$

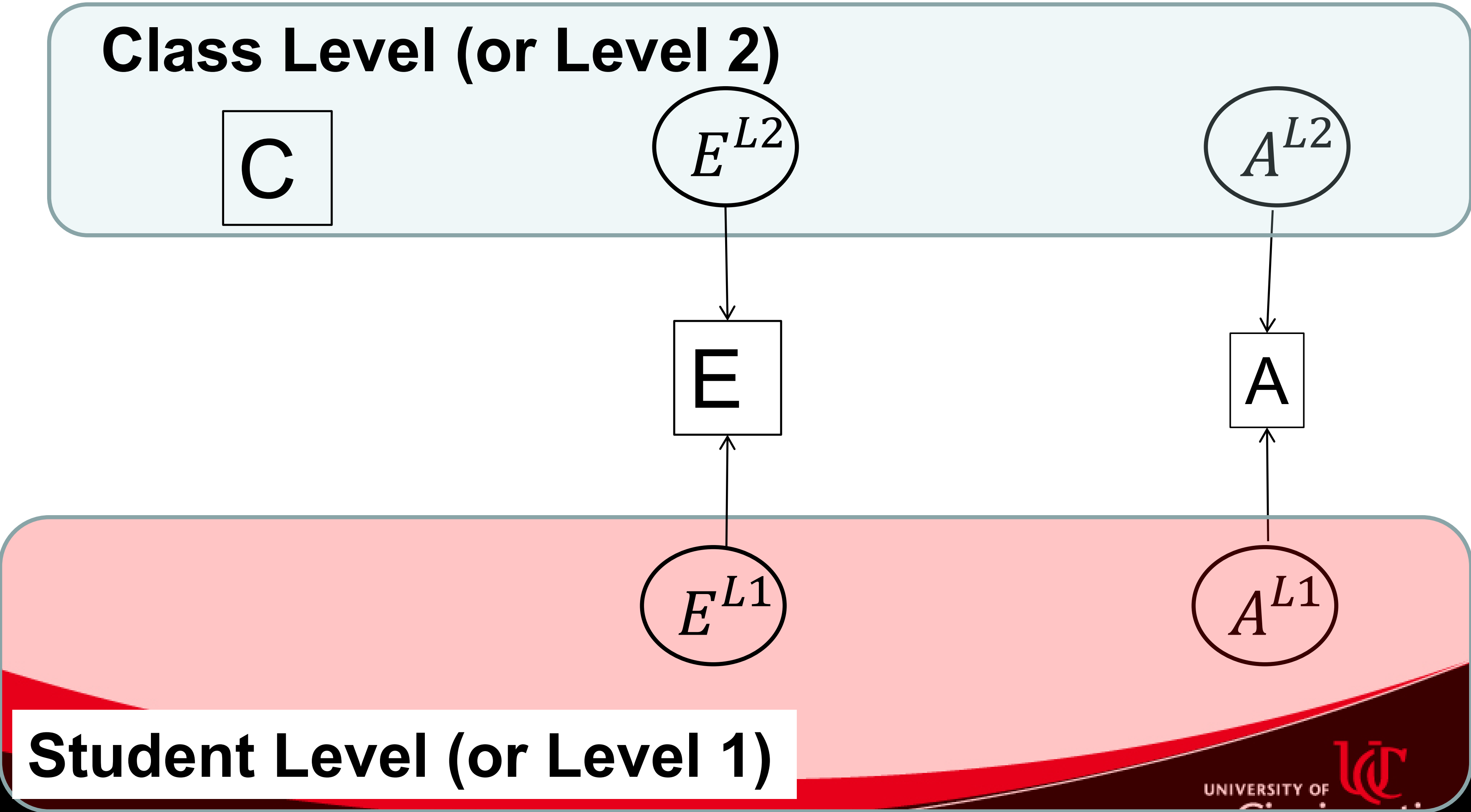
Example in R

Simple 2-1-1 Mediation

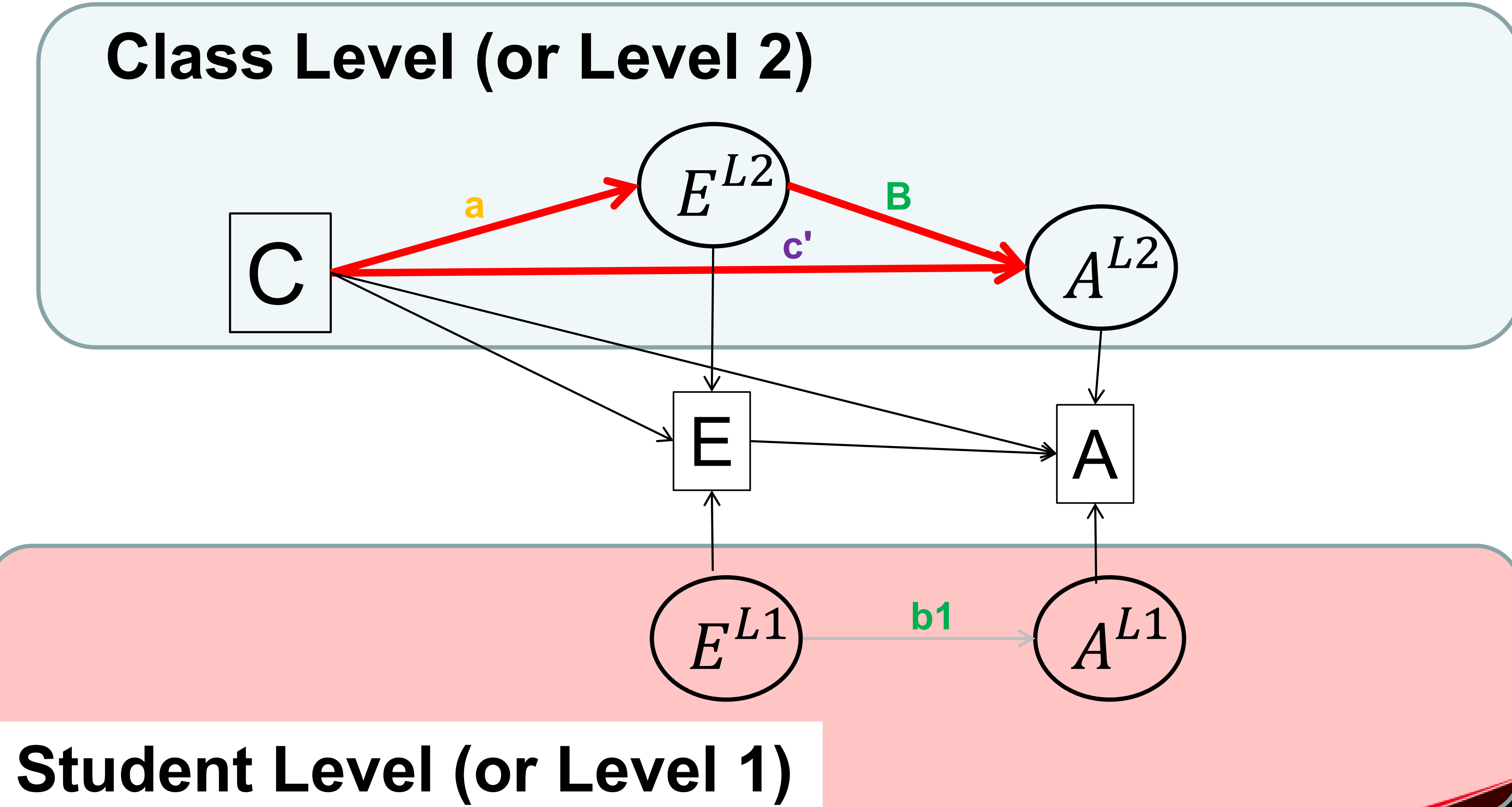
Imagine a cluster randomized trial that assigns teachers to use a new curriculum or the conventional curriculum. The theory of action underlying the curriculum is that it will improve student engagement which will in turn improve achievement.



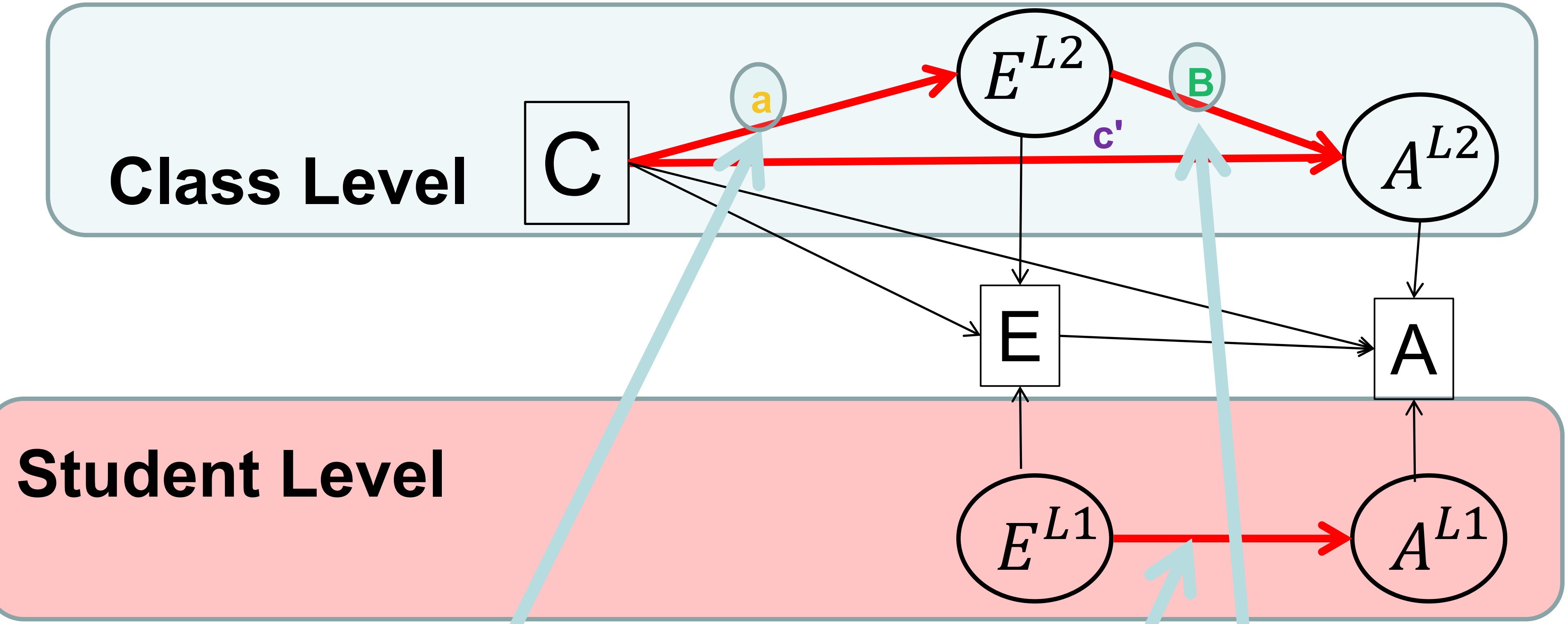
Multilevel Decomposition of Endogeneous Variables



Classroom Engagement



Classroom Engagement



Mediator Model

$$E_{ij} = \beta_{0j} + r_{ij},$$
$$\beta_{0j} = \gamma_{00} + aC_j + u_{0j}$$

Outcome Model

$$Y_{ij} = \beta_{0j} + b_1 E^{L1} + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + c'\ C_j + B\, E^{L2} + u_{0j}$$

Example in R

End of Session 20

- Break until 1030am

- Questions, Comments, & Feedback
  ben.kelcey@gmail.com

UNIVERSITY OF Cincinnati