Modeling Mediation: Causes, Markers, and Mechanisms

Stephen W. Raudenbush
University of Chicago

Address at the Society for Resesarch on Educational Effectiveness,Washington, DC, March 3, 2011. Many thanks to Guanglei Hong and Sean Reardon for sharing their ideas for this talk. Research reported here was supported by the W.T. Grant Foundation for a project entitled “Building Capacity to Support Group Randomized Experiments.”

Sea Change in Ed Research!

- Role of Experiments
  - Cluster randomized trials
  - Smart designs, Well powered!
- Sound Quasi-Experiments
  - Regression Discontinuity
  - Propensity score adjustment
  - Instrumental variables

What Experiments Can Tell Us

- Clear answer to sharp causal questions:
  - Average effect of assignment to treatments
  - If high compliance, average effect of an intervention
  - For whom does the intervention work?
  - In what settings does it work?

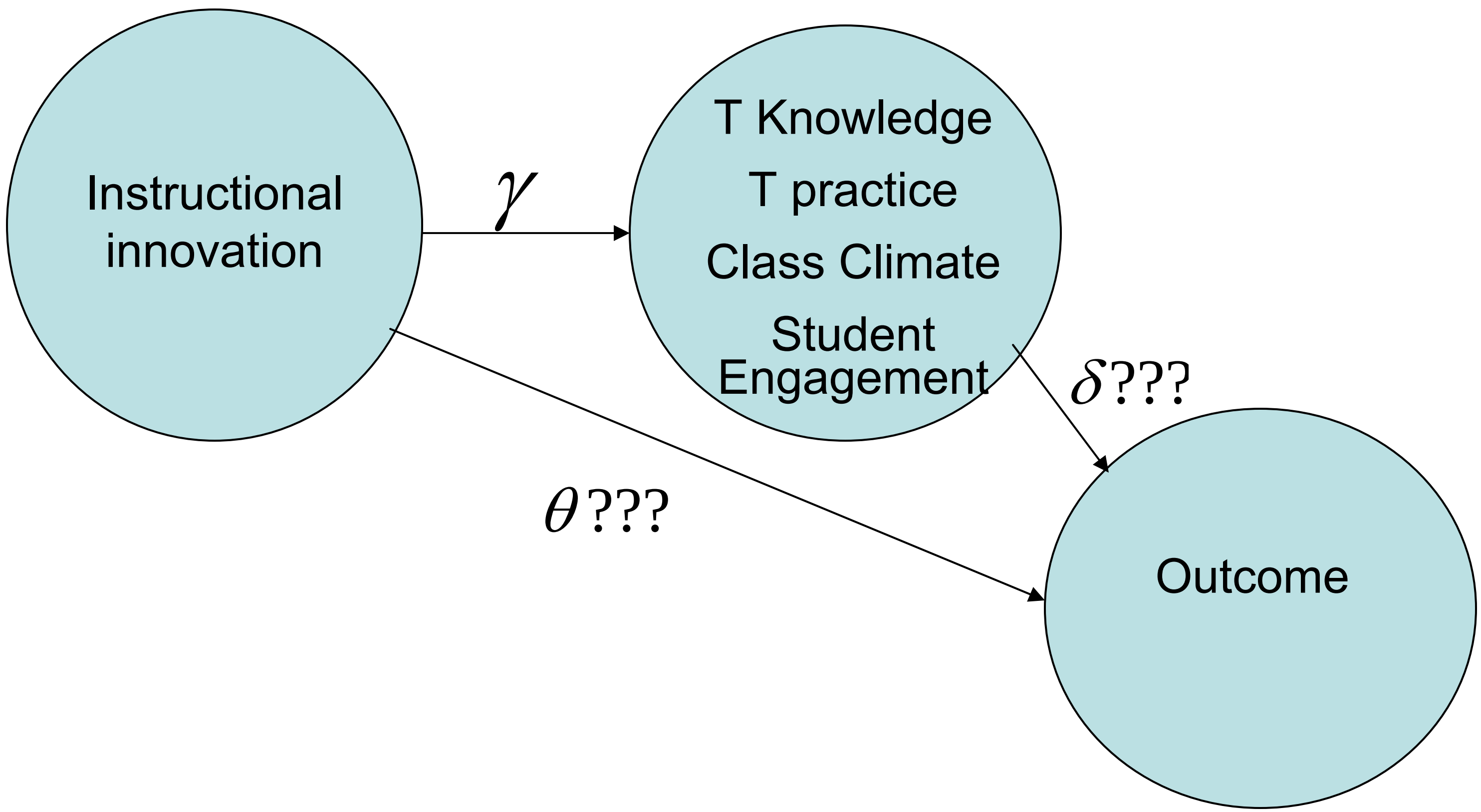
What about “How?” and “Why?”

- Example: IES study of pre-school curricula
  - Tells us which, on average, work better
  - But why
    - More time of academic instruction?
    - Better content?
    - Increase in teacher skill?
    - Compensate for lack of teacher skill?
    - Heighten child engagement?

Questions of Mediation

- In lab experiments, conduct a subsequent "factorial" randomized study
  - E.g., curricula by instructional time
- Very costly in large-scale field studies

Major Progress in Measurement



Some Major Progress

- We can ask
  - Was the treatment implemented
  - Did the treatment assignment affect practice?
  - Did the treatment increase student engagement

- These help interpret average causal effects

More challenging

- Did the treatment affect *Y* by
  - Increasing teacher knowledge?
  - Increasing teacher skill?
  - Improving class climate?
  - Increasing student engagement?
- These are questions of mediation

Models for Mediation

- Conventional methods under scrutiny
- Alternatives
  - Natural direct and indirect effects
  - Controlled direct effects
  - Direct Effects within Principal Strata
  - Instrumental variables for indirect effects
- Each requires assumptions
- Methods of Estimation not widely accessible

Punch Line

- Each approach answers a different question
- So let’s think hard about our questions

Organization of Talk

Consider questions about

- Demographic markers that predict access to treatments that predict outcomes
- Treatments that predict access to future treatments that predict outcomes
- Treatments that affect surrogate markers that predict outcomes

In light of these questions consider

- Oaxaca Decomposition
- Natural Direct Effects
- Controlled Direct Effects
- Principal Stratification
- Instrumental Variables

Reflect on the larger project: finding mechanisms

“All-purpose” mediation model

(Blau and Duncan, 1965; Baron and Kenny and Baron 1986)

Predictor	*Mediator*	*Outcome*
Demographic marker (parent SES)	Treatment (years of ed.)	Outcome (earnings)
Treatment (new curriculum v old)	Mediating treatment (quality of instruction)	Outcome (learning)
Treatment (pre-school program)	Surrogate marker (early vocab)	Outcome (Reading comp)

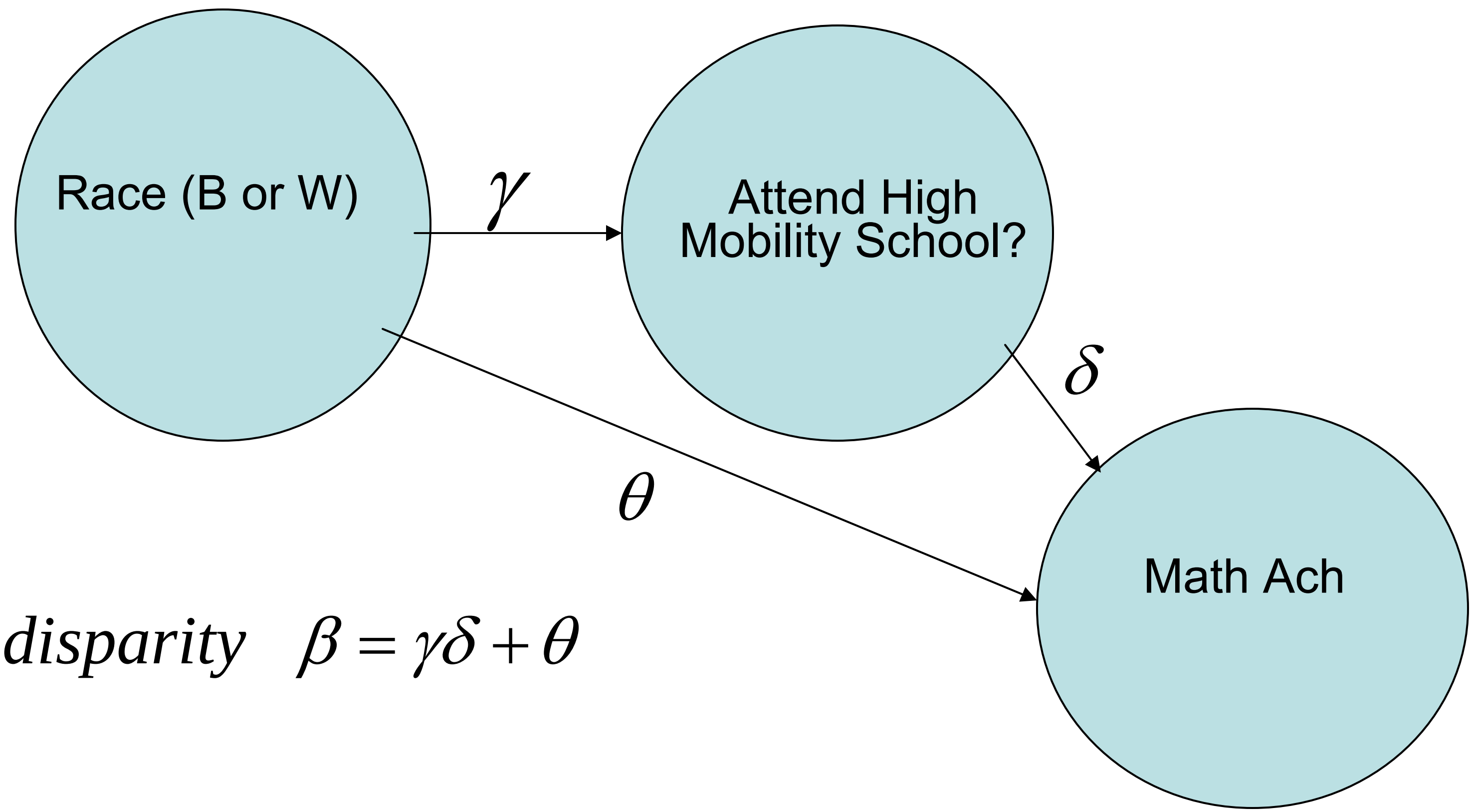
Start with Case 1

Demographic Marker-->Treat-->Y

Example:

Does school mobility explain B-W test score gap? (Raudenbush, Jean, Art, 2011)

Conventional Model



$$Racial\ disparity \quad \beta = \gamma\delta + \theta$$

But we found differential effects!

$$\gamma = (\phi_b - \phi_w) = \text{differential exposure}$$

$$\delta_b - \delta_w = \text{differential vulnerability}$$

$$\textit{Oaxaca decomposition}:$$

$$\begin{aligned}\beta = (\phi_b - \phi_w)\delta_b \qquad & \left(\textit{indirect path}\right) \\ + \theta_{\text{low}} + \phi_w(\delta_b - \delta_w) \quad & \left(\textit{direct path}\right)\end{aligned}$$

Decomposition is not unique

$$\beta = (\phi_b - \phi_w)\delta_b \quad \left(indirect\ \ path\right)$$
$$+\theta_{\mathrm{low}} + \phi_w(\delta_b - \delta_w) \quad \left(direct\ \ path\right)$$

$$= (\phi_b - \phi_w)\delta_w \quad \left(indirect\ \ path\right)$$
$$+\theta_{\mathrm{low}} + \phi_b(\delta_b - \delta_w) \quad \left(direct\ \ path\right)$$

Alternative expression

$$
\begin{aligned}
\beta = {} & (\phi_b - \phi_w)\delta_b \quad \left(indirect \;\; path\right) \\
& + \phi_w \theta_{\mathrm{high}} + (1 - \phi_w)\theta_{\mathrm{low}} \quad \left(direct \;\; path\right)
\end{aligned}
$$

Case 1 is easy!

- We have to estimate only 1 causal effect for Blacks and 1 for Whites
- Assume ignorable assignment to treatment within race
- Eg: How much does Class size reduction reduce the B-W gap (Krueger and Whitmore, 19xx)

Case 2

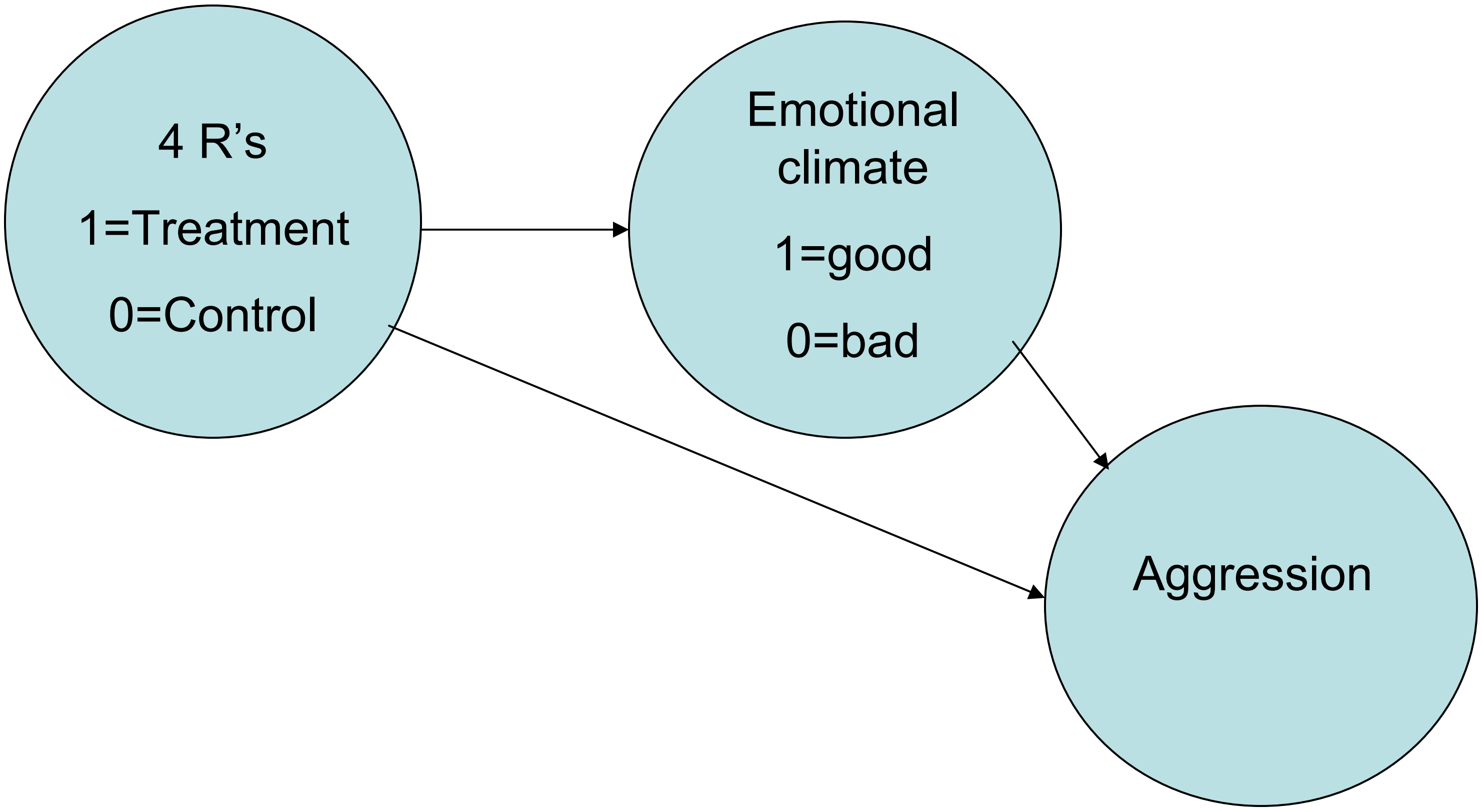
Treatment → Mediating Treatment → Y

Example:

Does a new curriculum reduce aggression by improving the emotional climate of the classroom?

(VanderWeele, Hong, Jones, and Brown, 2011)

Assignment to 4Rs Induces Better Climate

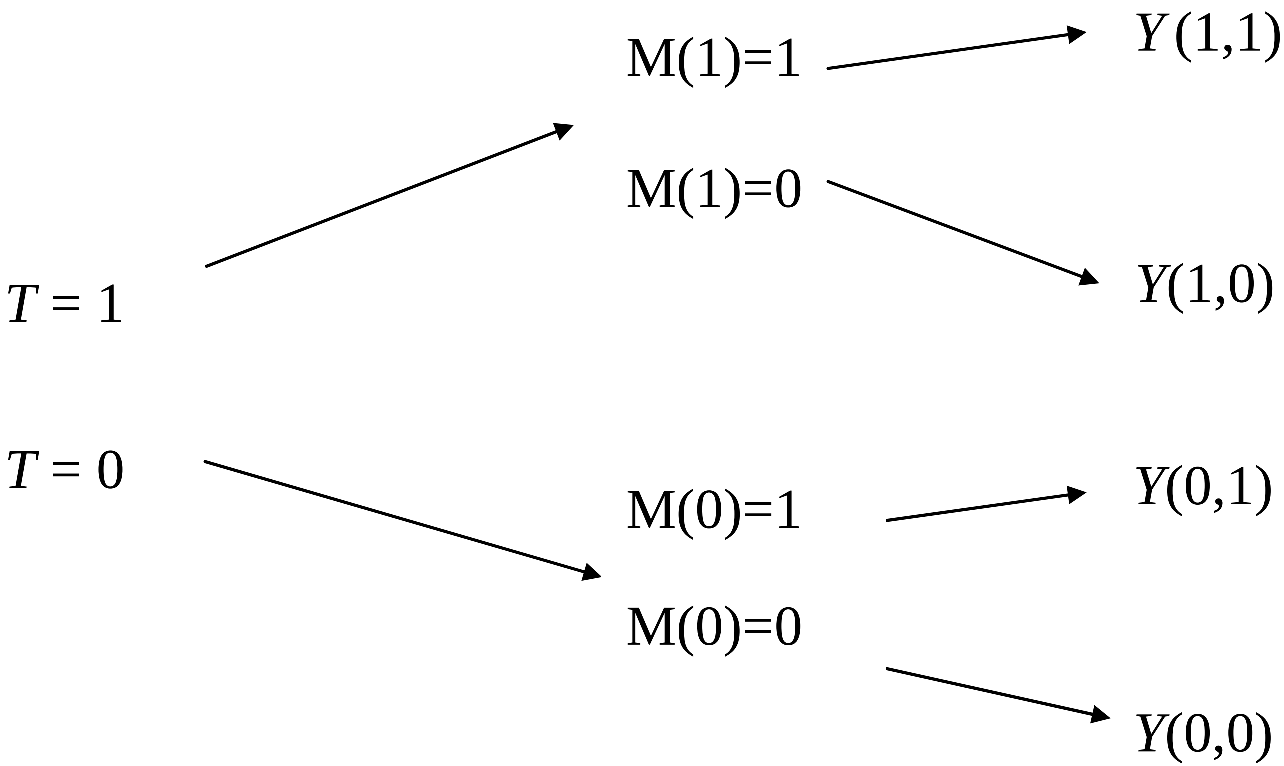


Must estimate causal effects of a sequence of two treatments!

The first (randomized) treatment increases the probability of exposure to the second (non-randomized) treatment.

Assumptions are subtle and strong!

Person-specific Potential Outcomes



Person-specific Causal Effects

- Effect of *T* on *M*

$$\Gamma = M(1) - M(0)$$

- Effect of *M* on *T*

$$\Delta_1 = Y(1,1) - Y(1,0) \quad \text{effect if } T = 1$$

$$\Delta_0 = Y(0,1) - Y(0,0) \quad \text{effect if } T = 0$$

- *Direct Effect of T on Y*

$$\Theta_1 = Y(1,1) - Y(0,1) \quad \text{effect if } M = 1$$

$$\Theta_0 = Y(1,0) - Y(0,0) \quad \text{effect if } M = 0$$

Natural Direct and Indirect Effects

- Total Effect: person-specific "Oaxaca Decomposition!"

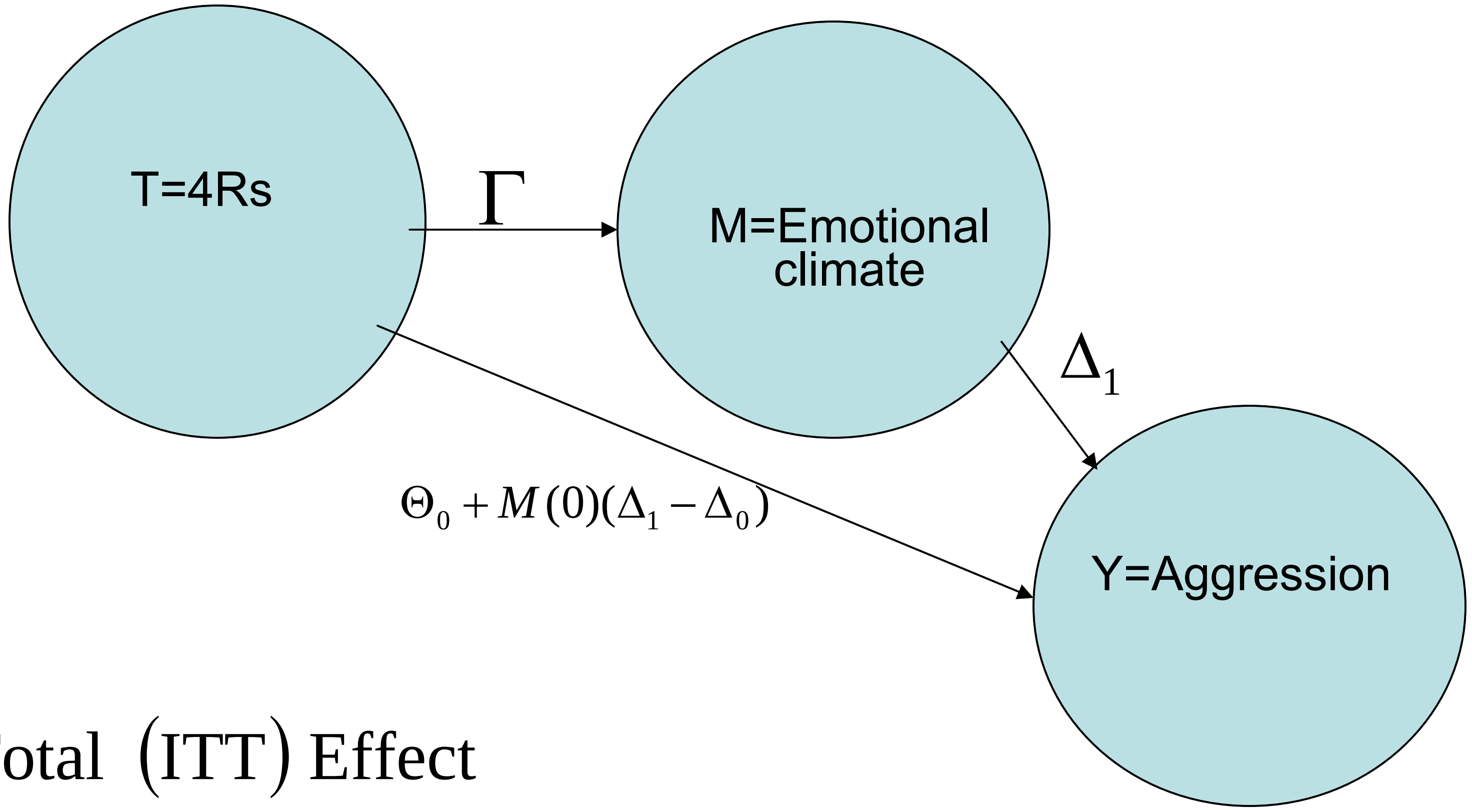
$$B = Y(1, M(1)) - Y(0, M(0))$$

$$\begin{aligned} &= Y(1, M(1)) - Y(1, M(0)) \quad \text{natural indirect effect} \\ &\quad + Y(1, M(0)) - Y(0, M(0)) \quad \text{natural direct effect} \end{aligned}$$

$$= \Gamma\Delta_1 \quad \textit{natural indirect effect}$$

$$+ \Theta_0 + M(0)(\Delta_1 - \Delta_0) \quad \textit{natural direct effect}$$

Person-specific Causal Model



$$\text{Total } (\text{ITT}) \text{ Effect}$$

$$B = \Gamma\Delta_1 + \Theta_0 + M(0)(\Delta_1 - \Delta_0)$$

But what is the *average* causal effect?

$$E(B) = \beta = E[\Gamma\Delta_1 + \Theta_0 + M(0)(\Delta_1 - \Delta_0)]$$

$$= \gamma\delta_1 + Cov(\Gamma, \Delta_1)$$
$$+ \Pr(M(0) = 1)(\delta_1 - \delta_0) + Cov(M(0), \Delta_1 - \Delta_0)$$

$$= \gamma\delta_1 + \theta_0 + \Pr(M(0) = 1)(\delta_1 - \delta_0)$$

(if we assume away covariances)

$$= \gamma\delta + \theta$$

(if we assume away differential effect of mediator)

Summary on Natural Direct and Indirect Effects

- Assume ignorable assignment to T
- Assume ignorable assignment to M within levels of T
- Assume away troublesome covariances
- Estimate via a regression method (Peterson et al.) or a non-parametric weighting method (Hong)
- Assume in addition no differential effects and linearity and use conventional methods!

Controlled Direct Effects

- Estimate

$$\theta_0 = E[Y(1,0) - Y(0,0)]$$
$$\theta_1 = E[Y(1,1) - Y(0,1)]$$

- Still need ignorable assignment to T
- Ignorable assignment to M given T
- Eliminates Troublesome covariances!
- But no estimation of indirect effect

Instrumental Variables

- Recall $B = \Gamma\Delta_1 \quad natural\ \ indirect\ \ effect$
  $+\Theta_0 + M(0)(\Delta_1 - \Delta_0) \quad natural\ \ direct\ \ effect$

- If we assume direct effect=0, we have

$$B = \Gamma\Delta \qquad \left(\Delta = Y(M(1) - Y(M(0))\right)$$

- Population-average effect

$$E(B) = \beta = E(\Gamma\Delta) = \gamma\delta + Cov(\Gamma, \Delta)$$

- Assume away covariance $\beta = E(\Gamma\Delta_1) = \gamma\delta_1$

- Assume $\gamma \neq 0$, we have $\beta / \gamma = \delta_1$

Summary on Instrumental Variables

- Assume ignorable assignment to *T*
- Assume no direct effect ("exclusion restriction")
- Assume away troublesome covariance
- Assume T affects M $(\gamma \neq 0)$

Case 3

Treatment → Surrogate marker →Y

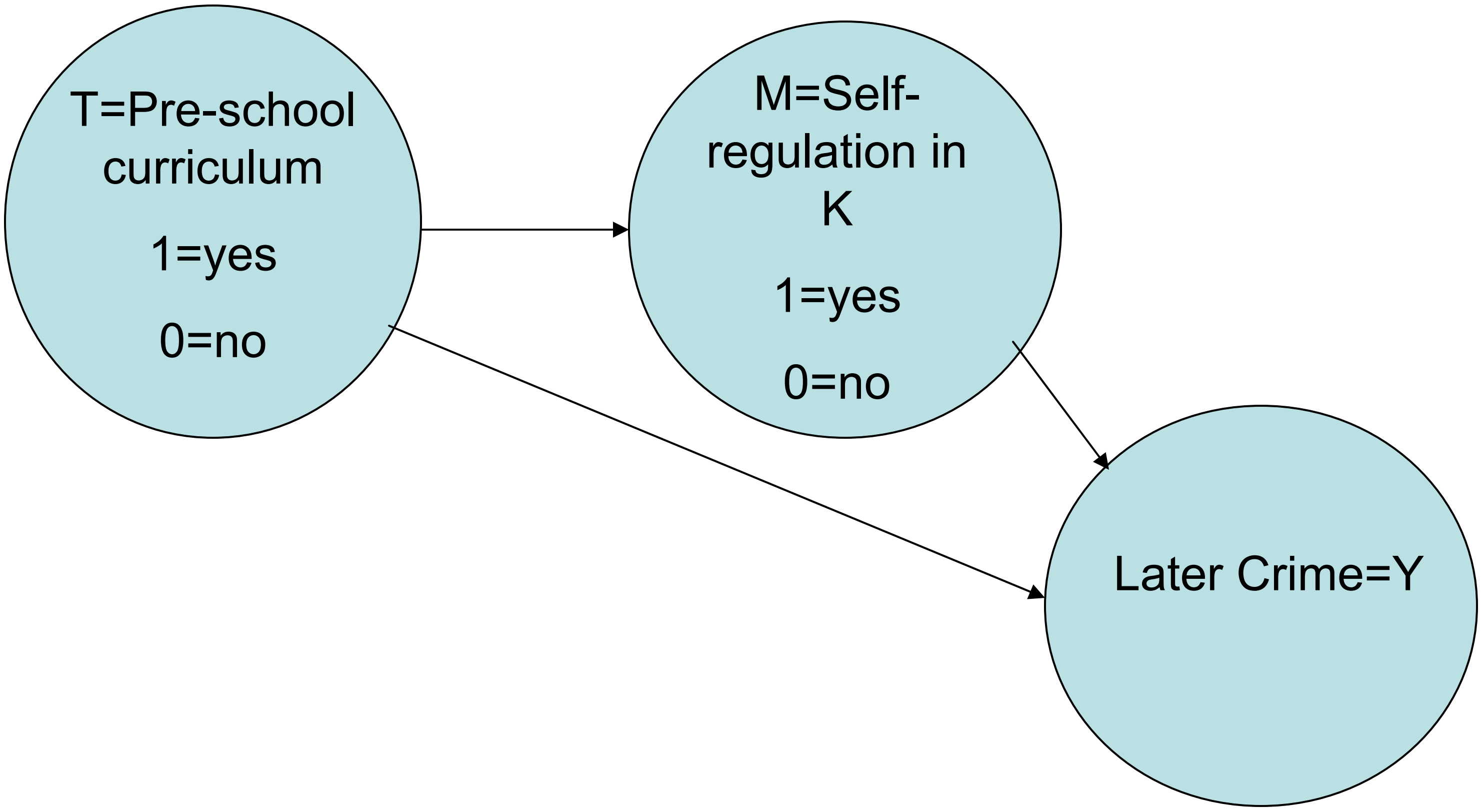
Example:

Does a new pre-school curriculum increase self regulation in K and hence reduce adolescent crime?

Mediator “stands in” for outcome=crime

Mediator is not a treatment

Case 3


T=Pre-school curriculum
1=yes
0=no
M=Self-regulation in K
1=yes
0=no
Later Crime=Y

Principal stratification

	M(1)	M(0)	Total Effect	Direct Effect
Compliers	1	0	Y(1,1)-Y(0,0)	----
Always-takers	1	1	Y(1,1)-Y(0,1)	θ_{AT}
Never-takers	0	0	Y(1,0)-Y(0,0)	θ_{NT}
Defiers	0	1	Y(1,0)-Y(0,1)	----

Summary on Principal stratification

- How a child will respond to intervention is a pre-treatment characteristic of that child
- Causal effects defined only within strata
- Direct effects defined for strata of children whose M is not changed by T.
- Strata are not observed, must be modeled
- No sharp point estimate of causal effects
- Seems most plausible when M is a surrogate marker rather than a second treatment

Summary

Case 1 (Demographic marker→Treat→Y)
- Use Oaxaca decomposition

Case 2 (Treat1→Treat 2→Y)
- Natural direct and indirect effects
- Controlled direct effects
- Instrumental variables

Case 3 (Treat→S. Marker→Y)
- Principal stratification

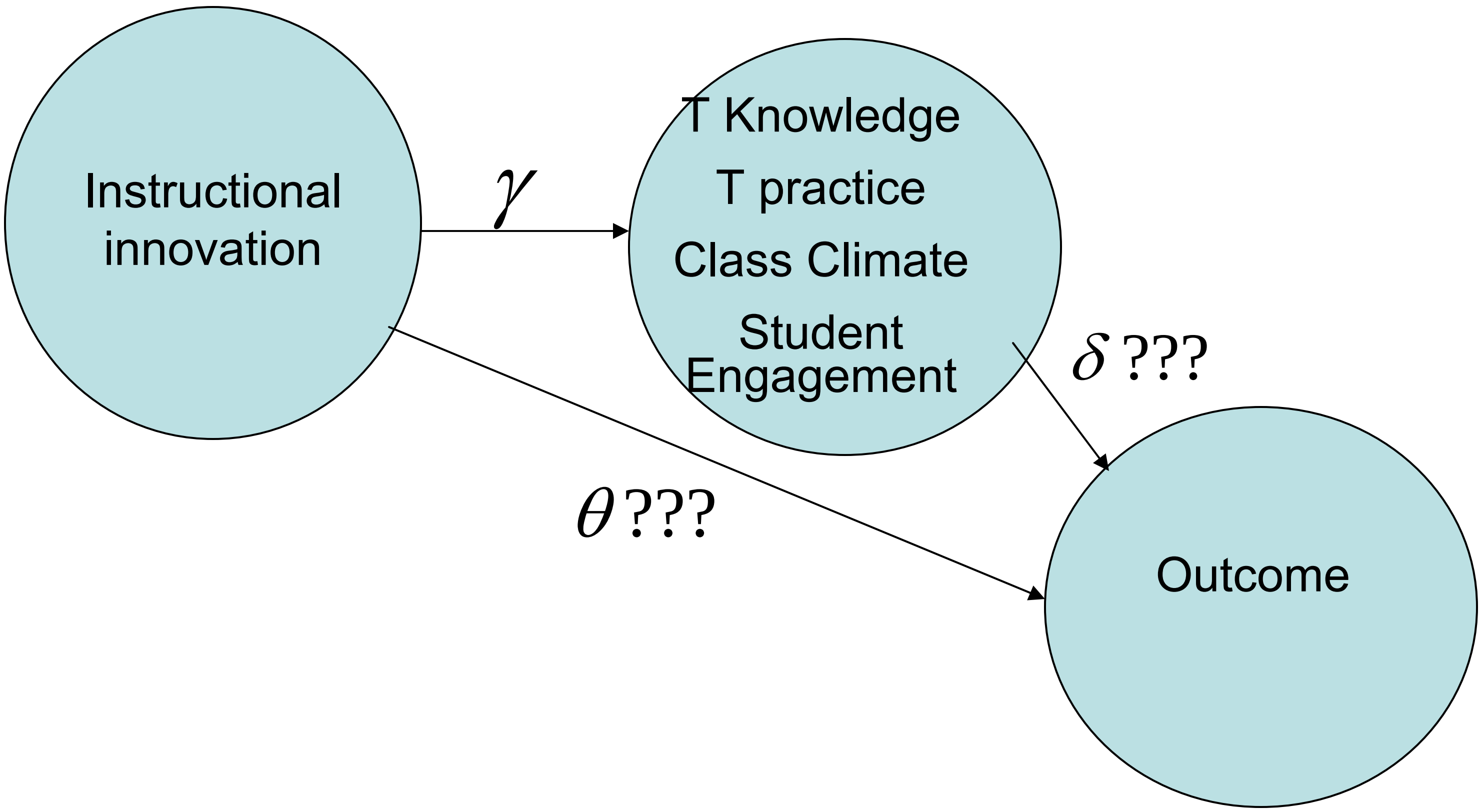
Conclusion

- Each method makes pretty strong assumptions
- The conventional "SEM" makes very strong assumptions
- We should pursue:
  - The method that best suits the question
  - That makes weakest assumptions
- We should state and evaluate the assumptions

However,…

- The project of discovering mechanisms and explaining effects requires a broad, multidisciplinary research agenda

Study “one arrow at a time”



References

Classical mediation

Baron, R.M., and Kenny, D.A., (1986). The moderator-mediator distinction in social psychology research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51,6,1173-1182.

Duncan, O.D. (1966). Path analysis: Sociological examples. *American Journal of Sociology*, 72,1-16.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, 5,3,161-215.

Oaxaca decomposition

Oxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14,3,693-709.

Natural and controlled direct effects

Hong, G. (2010). Ratio of mediator probability weighting for estimating natural direct and indirect effects. *Proceedings of the American Statistical Association Joint Statistical Meetings, Biometrics Section*, 2401-2415.

Pearl, J. (2001). Direct and indirect effects. *Proceedings of the American Statistical Association Joint Statistical Meetings*, 1572-1581.

Peterson, M.L.., Sinsi, S.E., van der Lan, M.G. (2006). Estimation of direct causal effects. *Epidemiology*, 17,3,276-284.

Instrumental variables

Angrist, J.D., and Krueger, A.B., (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, 15,4, 69-85.

Reardon, S.F. and Raudenbush, S.W. (2011). *Under What Assumptions do Site-by-Treatment Instruments Identify Average Causal Effects?* Stanford University School of Education.

Principal stratification

Frangakis, C.E., and Rubin, Principal stratification in causal inference. *Biometrics*, 58(1),21-29.

Page, L.C. (2010). *Comparing methodological approaches for conducting mediation analysis: A case for principal stratification*. Harvard University Graduate School of Education.

Empirical studies cited

Krueger, A.B., and Whitmore, D.M. (2001). The effect of attending a small class in the early grades on college test-taking and middle school test results: Evidence from Project STAR. *The Economic Journal*, 111,468, 1-28.

Raudenbush, S. W., Jean, Marshall, and Emily Art. (in press) Year-by-Year cumulative impacts of attending a high-mobility elementary school on children's mathematics achievement in Chicago, 1995 – 2005. In *Social Inequality and Educational Disadvantage*. Greg Duncan and Richard Murnane. (Eds.) Washington: Brookings.

VanderWeele, T.J., Hong, G., Jones, S., Brown, J. (2011). *Mediation and spillover effects in group-randomized trials: A case study of the 4Rs intervention*. Harvard University School of Public Health. Cambridge, MA.