Multilevel Mediation using SEM

Peter F. Halpin

NEW YORK UNIVERSITY

Outline

- Motivating example: OPEQ
- Conceptual review of (single-level) mediation: past and current approaches
- A crash course in (multilevel) SEM
- Example using Mplus

- Note: change of modeling framework (HLM to SEM), change of software (HLM to Mplus), change of example...

...Why all of these changes for mediation?

- Moderation: variation in program impact over pre-existing sub groups
  - adding interaction terms (new "X" variables)
  - also suggests changes to design of impact evaluations (e.g., multisite versus CRT; see Weiss, Bloom, & Brock, 2014)

...Why all of these changes for mediation?

- Moderation: variation in program impact over pre-existing sub groups
  - adding interaction terms (new "X" variables)
  - also suggests changes to design of impact evaluations (e.g., multisite versus CRT; see Weiss, Bloom, & Brock, 2014)
- Mediation: how does a program bring about changes in an outcome?
  - adding variables (mediators) between treatment and outcome (new "Y" variables)
  - requires a theory of change / logic model
  - multiple "Y" means multiple regression equations $\rightarrow$ SEM

The OPEQ Study[1]

- Study: **OP**portunities of **EQ**uitable access to basic education
  - undertaken in the Democratic Republic of the Congo (DRC) between 2011 and 2014
- Program: Learning in Healing Classrooms (LHC)
  - integrated reading and math curricula with a focus on socio emotional learning (SEL)
  - in-service teacher training and coaching via teacher learning circles

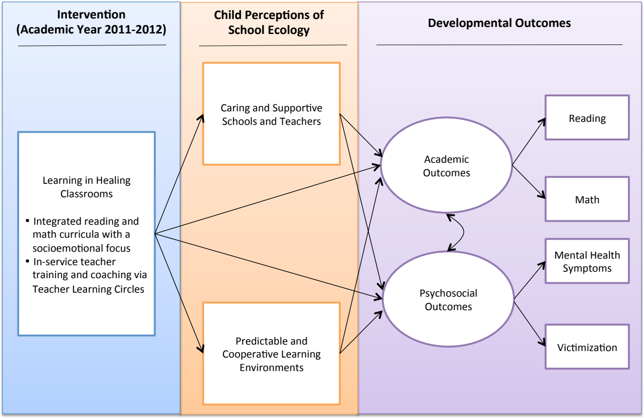
[1] OPEQ study was undertaken in the Democratic Republic of the Congo (CDR) between 2011 and 2014. The LHC program was developed by the International Rescue Committee in collaboration with the DRC Ministry of Education. The impact evaluation in DRC was implemented in partnership with RTI International, the Flemish Association for Development Cooperation and Technical Assistance, and the Institute of Human Development and Social Change at New York University, and was funded by USAID.

OPEQ: data

- Impact evaluation of LHC: focussed on literacy, numeracy, and SEL outcomes of children in grades 2-5
  - our data: "midline" (AY Feb - April 2012)
  - $K = 40$ clusters of schools; $J = 64$ schools; $N = 4,208$ students

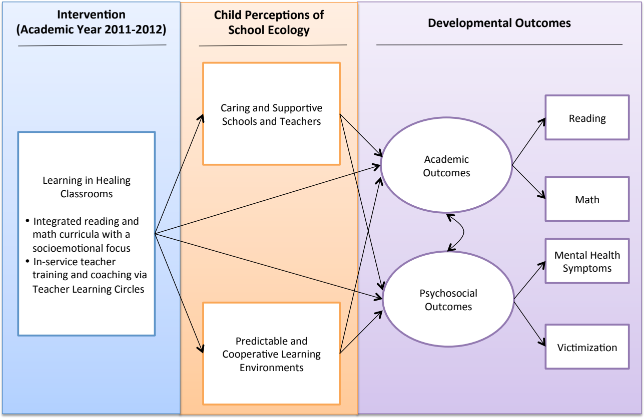
OPEQ: theory of change

Hypothesized Model of Influence of a School-Based Social-Emotional Learning Intervention on Children's Perceptions of their School Ecologies and Academic and Psychosocial Outcomes



OPEQ: theory of change

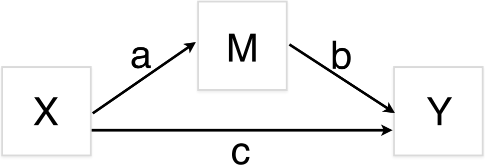
Hypothesized Model of Influence of a School-Based Social-Emotional Learning Intervention on Children's Perceptions of their School Ecologies and Academic and Psychosocial Outcomes



Our focus for now: LHC, Caring and Supportive Schools and Teachers, Math.

OPEQ: our example for today

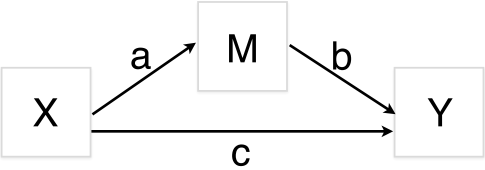
- "The Baron & Kenny (1986) triangle"



X = LHC (**treatment**)
M = Caring and supporting classrooms (**mediator**)
Y = Math (**outcome**)

OPEQ: our example for today

- "The Baron & Kenny (1986) triangle"



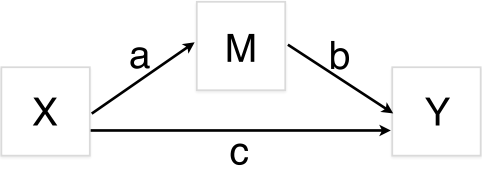
X = LHC (**treatment**)
M = Caring and supporting classrooms (**mediator**)
Y = Math (**outcome**)

a = **direct** effect of X on M
b = **direct** effect of M on Y
c = **direct** effect of X on Y (also denoted c')

ab = **indirect** effect of X on Y
ab + c = **total** effect of X on Y

total = **direct** + **indirect**

OPEQ: our example for today



- Interpreting effect sizes when $a, b, c > 0$:
  - $d = ab + c$ is the total effect; usual interpretation (e.g., treatment effect)
  - $d \geq ab$: the indirect effect cannot be larger than the total effect
  - ab/d is proportion of the total effect that is mediated
- When a, b, c have different signs – situation is not clear cut (e.g., direct and indirect effects, without total effects)

Why do we care about mediation?

Why do we care about mediation?

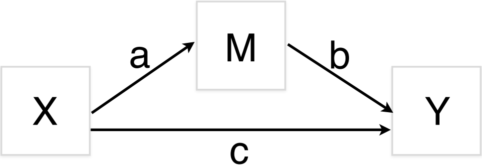
- To open "the black-box" of program effectiveness

Why do we care about mediation (Hong, 2012)?

1) To find out why an intervention failed to improve student outcomes:
    a) perhaps it did NOT have an impact on “targeted intermediate experiences” $\rightarrow$ problem with implementation
    b) perhaps it did have the expected impact on intermediate experiences $\rightarrow$ problems with theory
    c) perhaps the effect on the outcome was offset by worsening of impacts on the intermediate experiences $\rightarrow$ problems with theory

OPEQ: our example for today

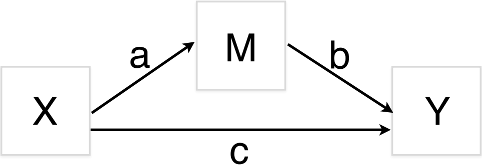
- What are case 1a) and 1b) in terms of the diagram?
  - assume all coefficients should be positive "in theory"



- No effect of program: ???
- Program did NOT effect intermediate experience: ???
- Program had expected effect on intermediate experience: ???
- Program had negative effect on intermediate experience: ???

OPEQ: our example for today

- What are case 1a) and 1b) in terms of the diagram?
  - assume all coefficients should be positive "in theory"



- No effect of program: $ab + c = 0$
- Program did NOT effect intermediate experience: $a = 0$
- Program had expected effect on intermediate experience: $a > 0$
- Program had negative effect on intermediate experience: $a < 0$

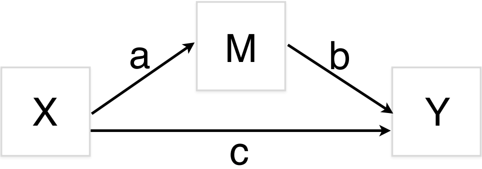
Why do we care about mediation (Hong, 2012)?

1) To find out why an intervention failed to improve student outcomes:
    a) perhaps it did NOT have an impact on "targeted intermediate experiences" $\rightarrow$ problem with implementation
    b) perhaps it did have the expected impact on intermediate experiences $\rightarrow$ problems with theory

2) Even if an intervention does improve student outcomes, the theory behind it may be wrong
    - does changing the intermediate experiences lead to a change in student outcomes?

OPEQ: our example for today

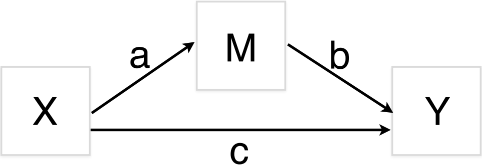
- What is case 2) in terms of the diagram?
  - assume all coefficients should be positive “in theory”



- Program did effect outcome: ???
- Changing the intermediate experience leads to change in the outcome: ???

OPEQ: our example for today

- What is case 2) in terms of the diagram?



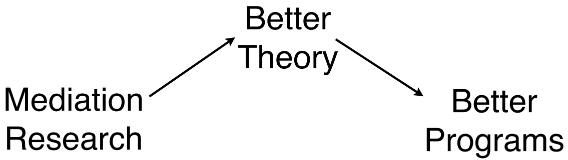
- Program did effect outcome: $ab + c \neq 0$
- Changing the intermediate experience leads to change in the outcome: $b = 0$

Why do we care about mediation?

- Your thoughts?
  - Are there additional reasons to consider mediation?
  - How would you describe the general purpose of mediation in program evaluation?
  - "CRTs are designed to definitively answer the question: Did the program work?" What does mediation add to this view of program evaluation?

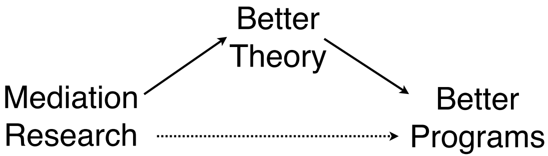
Why do we care about mediation?

- Studying mediation leads to better theory; better theory leads to better programs

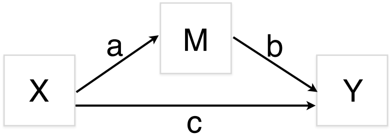


Why do we care about mediation?

- Studying mediation leads to better theory; better theory leads to better programs

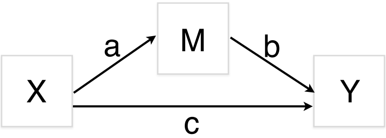


Past and current approaches to (single-level) mediation



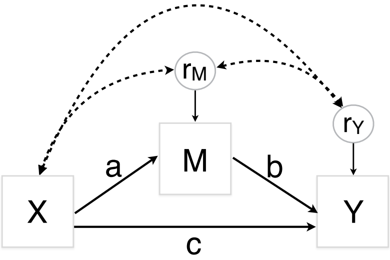
- Current approaches have considered defining causal mediation effects when:
  1) there is confounding / omitted variables
  2) there is an interaction between treatment and mediator
  3) linear regression doesn't apply

Past and current approaches to (single-level) mediation



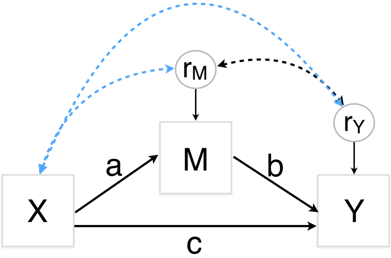
- Current approaches have considered defining causal mediation effects when:
  1) there is confounding / omitted variables
  2) there is an interaction between treatment and mediator
  3) linear regression doesn't apply
- This is not intended as a comprehensive review; see references

Current approaches: confounding



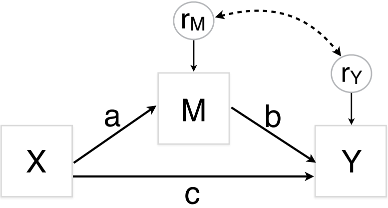
- Endogeneity:
  - the problem: "Variables are correlated with residuals"
  - one possible source of endogeneity: confounding / omitted variables
  - the gold standard solution for confounding: randomization

Current approaches: confounding



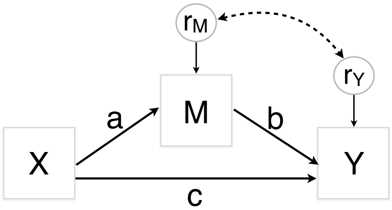
- Randomization of treatment "protects" the $X \rightarrow M$ and $X \rightarrow Y$ relationships...

Current approaches: confounding



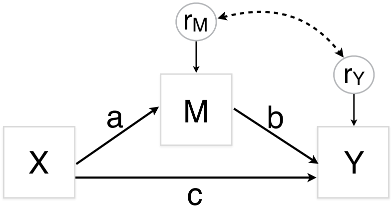
- ... but it doesn't protect the $M \to Y$ relationship

Current approaches: confounding



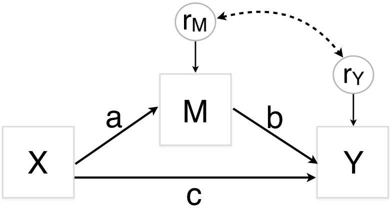
- ... but it doesn't protect the $M \rightarrow Y$ relationship
  - What are some plausible explanations of confounding here ?

Current approaches: confounding



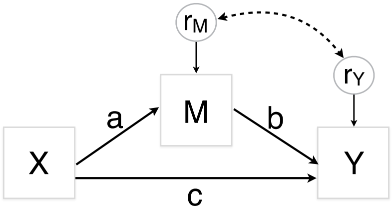
- Different approaches to causal mediation deal with this problem differently (see Page, 2012)
  - Principal stratification: estimate c assuming indirect effect $= 0$
  - Instrumental variables (2SLS): estimate indirect effect assuming $c = 0$
  - Regression based approaches: assume $\mathrm{cor}(r_M, r_Y) = 0$, but conduct sensitivity analysis (e.g., Imai, Keele, & Tingley, 2010)

Current approaches: confounding



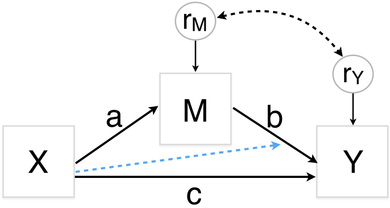
- More approaches:
  - SEM / graphical models: add more exogenous variables, apply rank and order rules (e.g., Bollen, 1989; see also Pearl, 2014)
  - Ratio-of-mediator-probability weighting (e.g., Hong & Nomi, 2012)
  - ....

Current approaches: more and more terminology



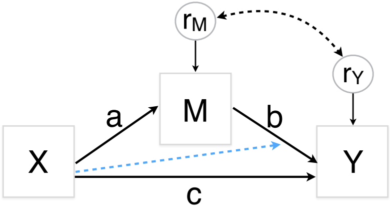
- Related problem: defining causal estimands for mediation (see Vanderweele & Vansteelandt, 2009)
  - controlled direct effects: set mediator to fixed value
  - natural direct and indirect effects: set mediator to value in counterfactual treatment condition
  - pure and total direct and indirect effects: how to parse out interactions?

Current approaches: treatment $\times$ mediator interaction



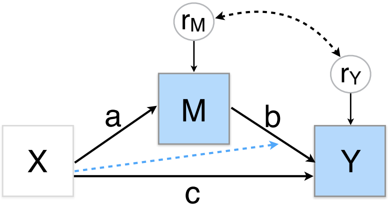
- The effect of the mediator on the outcome is not the same in the two treatment conditions
  - e.g., X = teacher instructional practices; M = student motivation; Y = student learning

Current approaches: treatment $\times$ mediator interaction



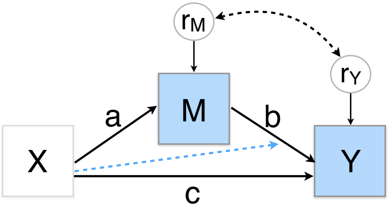
- The effect of the mediator on the outcome is not the same in the two treatment conditions
  - e.g., X = teacher instructional practices; M = student motivation; Y = student learning
  - ab is no longer a valid estimate of (natural) indirect effects! (see Imai et al. 2010; Valeri & Vanderweele, 2013; Muthén & Asparourhov, 2015)

Current approaches: non-linear models



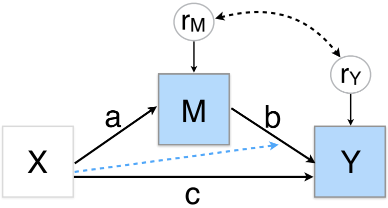
- What to do if M and / or Y are not continuous outcomes?
  - e.g., $Y =$ graduation ...

Current approaches: non-linear models



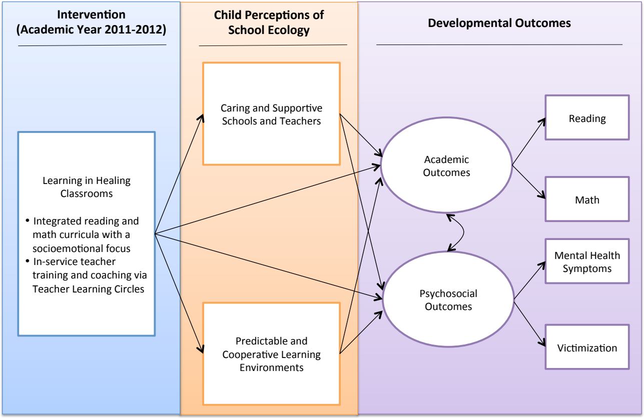
- What to do if M and / or Y are not continuous outcomes?
  - e.g., $Y =$ graduation ...
  - again: ab is no longer a valid estimate of (natural) indirect effects! (see Imai et al. 2010; Valeri & Vanderweele, 2013; Muthén & Asparourhov, 2015)

Current approaches: the new BKT



What do we need to get from BKT to OPEQ?

Hypothesized Model of Influence of a School-Based Social-Emotional Learning Intervention on Children's Perceptions of their School Ecologies and Academic and Psychosocial Outcomes



What do we need to get from BKT to OPEQ?

1) Multilevel context
2) Multiple mediators
3) Latent variables

What do we need to get from BKT to OPEQ?

1) Multilevel context
2) Multiple mediators
3) Latent variables

- Software options:
  - Currently available "boutique" software for causal mediation (e.g., Imai, et al. 2010; Valeri & Vanderweele, 2013)
    - in Stata: `paramed`; `ldecomp`; `medeff`; `gformula`

What do we need to get from BKT to OPEQ?

1) Multilevel context
2) Multiple mediators
3) Latent variables

- Software options:
  - Currently available "boutique" software for causal mediation (e.g., Imai, et al. 2010; Valeri & Vanderweele, 2013)
    - in Stata: `paramed`; `ldecomp`; `medeff`; `gformula`
  - SEM deals with 1) - 3), as well as many of the issues in raised by the new BKT (Muthén & Asparourhov, 2015)

What do we need to get from BKT to OPEQ?

1) Multilevel context
2) Multiple mediators
3) Latent variables

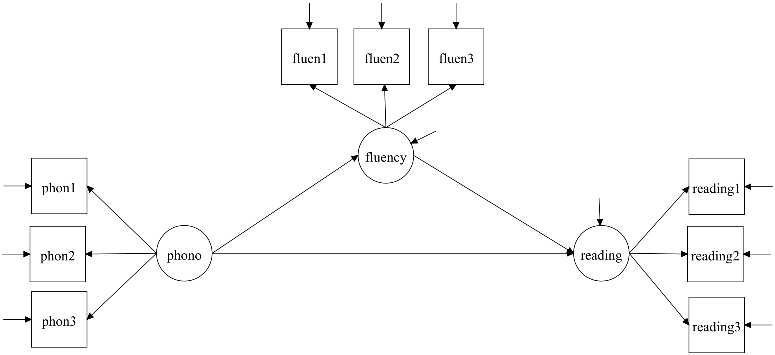
- Software options:
  - Currently available "boutique" software for causal mediation (e.g., Imai, et al. 2010; Valeri & Vanderweele, 2013)
    - in Stata: `paramed`; `ldecomp`; `medeff`; `gformula`
  - SEM deals with 1) - 3), as well as many of the issues in raised by the new BKT (Muthén & Asparourhov, 2015)
- A final consideration: statistical versus causal mediation

A crash course in (multilevel) SEM

The many layers of SEM

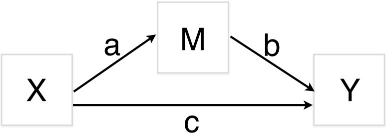
- Layer 1: The path diagram
  - Intuitive and therefore useful for model specification
  - But also easy to abuse
- Layer 2: Linear models
  - There are rules for translating a path diagram into a linear model
  - This is why drawing the diagram counts as model specification
- Layer 3: Covariance structures
  - Linear models also imply a model for the observed covariance matrix
  - Today this is important because multilevel models imply multiple covariance matrices
- Layer 4: Software
  - There are lots of software programs for SEM
  - MSEM: Mplus, GLLAMM (Stata), LISREL

Path diagram 1: with a measurement model

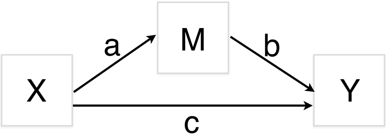


Path diagram 2: without a measurement model (path analysis)

- BKT is a path diagram!



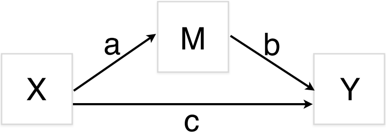
What about levels?



OPEQ example

Variable	Example	Level
X	LHC	???
M	Caring and Supportive Classrooms	???
Y	Math	???

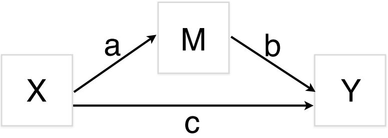
What about levels?



OPEQ example: "(2-2-1) design"

Variable	Example	Level
X	LHC	2
M	Caring and Supportive Classrooms	2
Y	Math	1

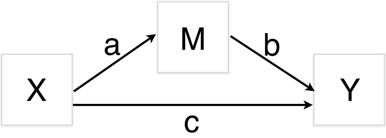
What about levels?



Teacher example:

Variable	Example	Level
X	Instructional Practices	???
M	Student motivation	???
Y	Learning outcomes	???

What about levels?



Teacher example: "(2-1-1) design"

Variable	Example	Level
X	Instructional Practices	2
M	Student motivation	1
Y	Learning outcomes	1

Setting up multilevel path diagrams

- Three types of variables in MSEM
  1) Variables that only vary at level 2 – group/cluster level variables
  2) Variables that only vary at level 1 – individual level variables with negligible ICC ($< .05$)
  3) Variables that vary at both levels

Setting up multilevel path diagrams

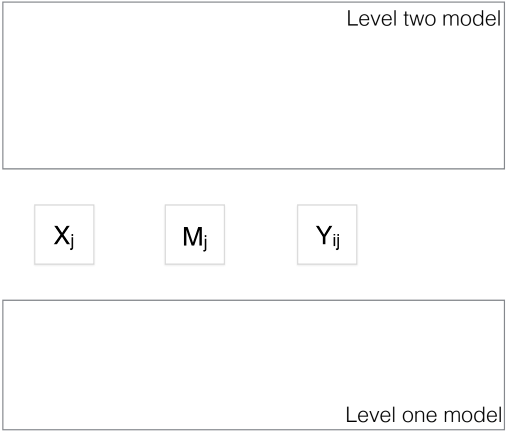
- Three types of variables in MSEM
  1) Variables that only vary at level 2 – group/cluster level variables
  2) Variables that only vary at level 1 – individual level variables with negligible ICC ($< .05$)
  3) Variables that vary at both levels

- Type 3) variables are always split into two parts:
  - 3a) A part that varies at level 2 only (group/cluster means)
  - 3b) A part that varies at level 1 only (group mean centered)

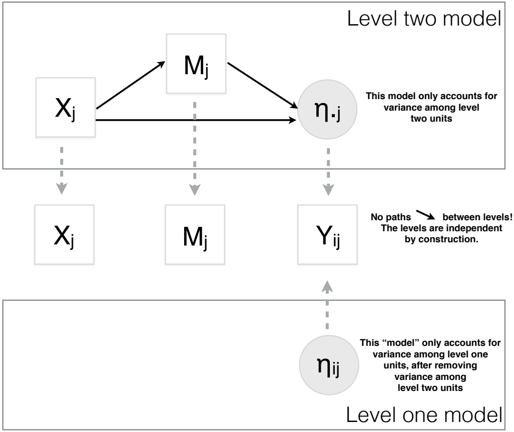
Setting up multilevel path diagrams

- Comparison with HLM
  - Level 1 variables are ALWAYS group mean centered (unless they have no variance at level 2)
  - Main difference: group means for ALL variables are treated as population parameters to be estimated (see Ludtke et al. 2008)
  - This is like treating all level 1 variables as having random intercepts, not just the Y variable

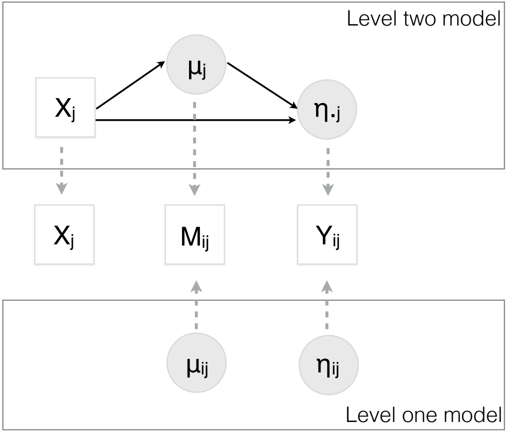
OPEQ example: two-level path diagram (2-2-1)



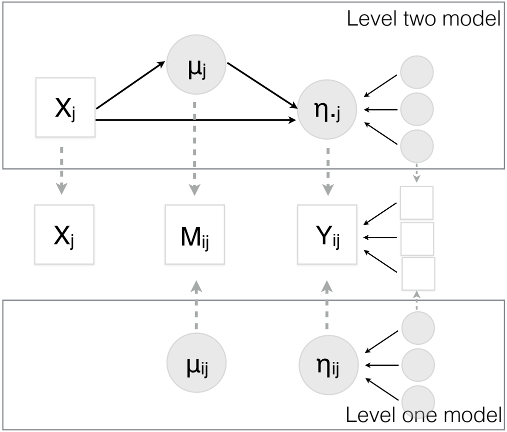
OPEQ example: two-level path diagram (2-2-1)



Teacher example again: (2-1-1)



Teacher example again: (2-1-1) with covariates



Layer 1 – Summary

- In multilevel (two-level) SEM, the basic idea is that we get two path diagrams

Layer 1 – Summary

- In multilevel (two-level) SEM, the basic idea is that we get two path diagrams
- The level 2 model explains variance over level 2 units (e.g., schools, classrooms) and includes:
  - All level 2 variables
  - The level 2 components (i.e., group means) of level 1 variables – but note dummy variables

Layer 1 – Summary

- In multilevel (two-level) SEM, the basic idea is that we get two path diagrams
- The level 2 model explains variance over level 2 units (e.g., schools, classrooms) and includes:
  - All level 2 variables
  - The level 2 components (i.e., group means) of level 1 variables – but note dummy variables
- The level 1 model explains variance over level 1 units (e.g., students), after removing and variance over level 2 units and includes:
  - The level 1 components (i.e., group mean deviations) of level 1 variables.

Layer 1 – Summary

- In multilevel (two-level) SEM, the basic idea is that we get two path diagrams

- The level 2 model explains variance over level 2 units (e.g., schools, classrooms) and includes:
  - All level 2 variables
  - The level 2 components (i.e., group means) of level 1 variables – but note dummy variables

- The level 1 model explains variance over level 1 units (e.g., students), after removing and variance over level 2 units and includes:
  - The level 1 components (i.e., group mean deviations) of level 1 variables.

- **The two levels are independent by construction.**

Layer 1 – Summary

- In and SEM context, multilevel mediation is somewhat of a misnomer:
  - There can be mediation at level 2; there can be mediation at level 1; but there is no mediation from level 2 to level 1 or vice versa
  - Mediation that involves a level-2 variable is always at level 2

The many layers of SEM

- Layer 1: The path diagram.
  - Intuitive and therefore useful for model specification
  - But also easy to abuse.
- Layer 2: Linear models
  - There are rules for translating a path diagram into a linear model
  - This is why drawing the diagram counts as model specification
- Layer 3: Covariance structures
  - Linear models also imply a model for the observed covariance matrix
  - Today this is important because multilevel models imply multiple covariance matrices
- Layer 4: Software
  - There are lots of software programs: today Mplus

Layer 2: model specification with linear equations

- Step 1: Treat the ALL level 1 (with ICC $> .05$) variables as the sum of two independent normally distributed variables
- We have seen this trick before (slide 9 and 10 of MLM refresher):
  - level 1 model: $Y_{ij} = \beta_{0j} + r_{ij}; \quad r_{ij} \sim N(0, \sigma^2)$
  - level 2 model: $\beta_{0j} = \gamma_{00} + u_{0j}; \quad u_{0j} \sim N(0, \tau_{00})$
  - mixed model: $Y_{ij} = \gamma_{00} + u_{0j} + r_{ij}$
- Change notation to incorporate same trick for other level 1 variables:
  - $Y_{ij} = \alpha_Y + \eta_{.j} + \eta_{ij}; \quad \eta_{.j} \sim N(0, \sigma^2_{\eta_{.j}}); \quad \eta_{ij} \sim N(0, \sigma^2_{\eta_{ij}})$

Layer 2: model specification with linear equations

- Step 2: Assume that all level 2 variables are jointly distributed, e.g.,

$$\begin{bmatrix} \eta_{.j} \\ X_j \\ M_j \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \alpha_X \\ \alpha_M \end{bmatrix}, \quad \begin{bmatrix} \sigma^2_{\eta_{.j}} & - & - \\ \sigma_{\eta_{.j}X_j} & \sigma^2_{X_j} & - \\ \sigma_{\eta_{.j}M_j} & \sigma_{X_jM_j} & \sigma^2_{M_j} \end{bmatrix} \right)$$

- Similar for level 1

Layer 2: model specification with linear equations

- Step 3: Represent the level 2 path diagram via linear equations, e.g.,
  - Model for outcome: $\eta_{.j} = \alpha_{\eta_{.j}} + cX_j + bM_j + u_j$
  - Model for mediator: $M_j = \alpha_{M_j} + aX_j + v_j$
  - Combined model: $\eta_{.j} = \alpha_{\eta_{.j}} + (c + ab)X_j + b[\alpha_{M_j} + v_j] + u_j$

- Similar for level 1
- Compare to omitted variable bias set up in OLS regression
  - M is the omitted variable, c is “causal effect” of X, and ab is the bias

Layer 3: Estimation, etc. via covariance matrices

- At each level, the linear equations imply a model for the covariance matrix of the variables, e.g.,

$$\mathrm{var}(\eta_{.j}) = (c + ab)^2\sigma^2_{X_j} + b^2\sigma_{u_j} + \sigma^2_{v_j}$$

- In practice, your SEM software handles this part
- Estimation equations and algorithms, goodness of fit, and identification are all via this representation of the model
- Note: goodness of fit is not applicable for saturated mediation models

Layer 4: Sofware

- See annotated script at the end of these slides for reference; also Preacher et al. 2011 reference in readings
- Will go over examples in lab time

General comments on MSEM in practice

- Is there level-2 variation in the outcome(s)?
  - Can get a "rough estimate" of the ICC for level-1 variables with 1-way ANOVA using the level-2 grouping variable ($SS_{between}/SS_{total}$).
  - If there is no level-2 variation in the outcome then there isn't anything to model at level 2.
  - If level-1 variables have low ICCs this can cause convergence problems (these variables are treated as "level 1 only")

General comments on MSEM in practice

- Are the variables of interest correlated at level 2?
  - Can get a “rough estimate” of the level-2 correlation matrix by aggregating level-2 variables to group means, compute correlation matrix

General comments on MSEM in practice

- Power: Are there enough level-2 units?
  - Power analysis for multilevel mediation is not like that for OD (exact distributions)
  - Asymptotic distributions for ab effects (e.g., Sobel test) known to perform poorly in finite samples (Shrout & Bolger, 2002)
  - Usual approach in SEM is power by simulation studies
  - But Li & Beretvas (2013) show power is $< .4$ for small effects ($ab = .09$) with 80 schools (also see Krull & MacKinnon 1999)

Questions or other topics?

Mplus: input file for (2-2-1)

```
TITLE:  2-2-1 example;

DATA:   FILE = mydata.csv;

VARIABLE:
    ! These are the variables IN THE DATASET
    NAMES = School_ID Y M X;

        ! These are the varibles you want to use;
    USEVARIABLES = School_ID Y M X;

    !This is the the cluster variable (can be 3-level)
    CLUSTER = School_ID;

    ! Variables at level 2 only
    BETWEEN = X M;

    ! Variables at level 1 only
    WITHIN = ;

    !Missing data code(s) used
    MISSING = ALL(999) ;

ANALYSIS:

    TYPE = TWOLEVEL;
    ! See Mplus manual for more options

MODEL:

    %WITHIN% !Level 1 model

     Y;

    %BETWEEN% !Level 2 mode

    Y ON M X;
    M ON X;

MODEL INDIRECT:

    ! Test indirect and total effects
    Y IND X;
```

Mplus general points: data

0) All data should be in the same file!
1) Data should be .csv or tab delimited (.dat) format (save in this format before using Mpus)
2) All entries in the data set must be numeric! This includes NA values (no blanks!)
3) no variable names can appear in the data set (see point 2) – names appear in the `NAMES` command of Mplus (see .input file)

Mplus general points: syntax

- commands must end with `:` and statements must end with `;`
- `ON` means "regression." Usage: `Y ON X1 X2 X3;`
- `USE VARIABLES` tells MPlus what variables you want to use, but `VARIABLES` tells Mplus what variables are in the data set: don't change the latter!
- use `MODEL INDIRECT:` command to get tests of total and indirect effects
- Questions about syntax: The Mplus user's guide
- Questions about analyses: The Mplus discussion board

References: Multilevel mediation

Krull, J. L., & MacKinnon, D. P. (1999). Multilevel mediation modeling in group-based intervention studies. *Evaluation Review*, *23*(4), 418–444. doi:10.1177/0193841X9902300404

Li, X., & Beretvas, S. N. (2013). Sample Size Limits for Estimating Upper Level Mediation Models Using Multilevel SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(2), 241–264. doi:10.1080/10705511.2013.769391

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. O. (2008). The multilevel latent covariate model: a new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, *13*(3), 203–29. doi:10.1037/a0012869

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2011). Alternative Methods for Assessing Mediation in Multilevel Data: The Advantages of Multilevel SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(2), 161–182. doi:10.1080/10705511.2011.557329

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, *15*(3), 209–33. doi:10.1037/a0020141

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing Multilevel Mediation Using Hierarchical Linear Models: Problems and Solutions. *Organization Research Methods*, *12*(4), 695–719. doi:10.5465/AMBPP.2008.33716518

References: Causal mediation

Hong, G. (2012). Editorial Comments. *Journal of Research on Educational Effectiveness*, *5*(3), 213–214. doi:10.1080/19345747.2012.688403

Hong, G., & Nomi, T. (2012). Weighting Methods for Assessing Policy Effects Mediated by Peer Change. *Journal of Research on Educational Effectiveness*, *5*(3), 261–289. doi:10.1080/15348431.2012.688421

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–34. doi:10.1037/a0020761

Muthén, B. O., & Asparouhov, T. (2015). Causal Effects in Mediation Modeling: An Introduction with Applications to Latent Variables. *Structural Equation Modeling*, *22*, 12–23. doi:10.1080/10705511.2014.935843

Page, L. C. (2012). Principal Stratification as a Framework for Investigating Mediational Processes in Experimental Settings. *Journal of Research on Educational Effectiveness*, *5*(3), 215–244. doi:10.1080/19345747.2012.688410

Pearl, J. (2014). Interpretation and Identification of Causal Mediation. *Psychological Methods*, *19*(4), 459–481. doi:10.1037/a0036434

Raudenbush, S. W., Reardon, S. F., & Nomi, T. (2012). Statistical Analysis for Multisite Trials Using Instrumental Variables With Random Coefficients. *Journal of Research on Educational Effectiveness*, *5*(3), 303–332. doi:10.1080/19345747.2012.689610

Ten Have, T. R., & Joffe, M. M. (2012). A review of causal estimation of effects in mediation analyses. *Statistical Methods in Medical Research*, *21*(1), 77–107. doi:10.1177/0962280210391076

Valeri, L., & Vanderweele, T. J. (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*, *18*(2), 137–50. doi:10.1037/a0031034

VanderWeele, T. J. (2012). Comments: Should Principal Stratification Be Used to Study Mediational Processes? *Journal of Research on Educational Effectiveness*, *5*(3), 245–249. doi:10.1080/19345747.2012.688412

Vanderweele, T. J., & Vansteelandt, S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and Its Interface*, *2*, 457–468. doi:10.4310/SII.2009.v2.n4.a7

References: Software

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–34. doi:10.1037/a0020761

Preacher, K. J. (2011). Multilevel SEM Strategies for Evaluating Mediation in Three-Level Data. *Multivariate Behavioral Research*, *46*(August 2014), 691–731. doi:10.1080/00273171.2011.589280

Valeri, L., & Vanderweele, T. J. (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*, *18*(2), 137–50. doi:10.1037/a0031034