New Methods for Studying Mediating Mechanisms in Developmental and Intervention Studies of Child Maltreatment

David P. MacKinnon

& Linda J. Luecken

Arizona State University

Presented at the 2021 Conference, "Innovative Methods in Child Maltreatment Research: The Latest Approaches for Improving Detection and Intervention"

September 2, 2021, University Park, Pennsylvania

***This research was supported in part by the National Institute on Drug Abuse (R37DA09757) and National Institute of Mental Health (MH083173).**

Outline

1. Background and Motivations
2. Statistical Testing and Confidence Limit Estimation
3. Inconsistent Mediation Models
4. Longitudinal Mediation Models
5. Measurement and Mediation
6. Causal Inference for Mediation

Website:

http://www.public.asu.edu/~davidpm/

Book:

MacKinnon, D. P. (2008; Second Edition) Introduction to Statistical Mediation Analysis. Mahwah, NJ: Erlbaum.

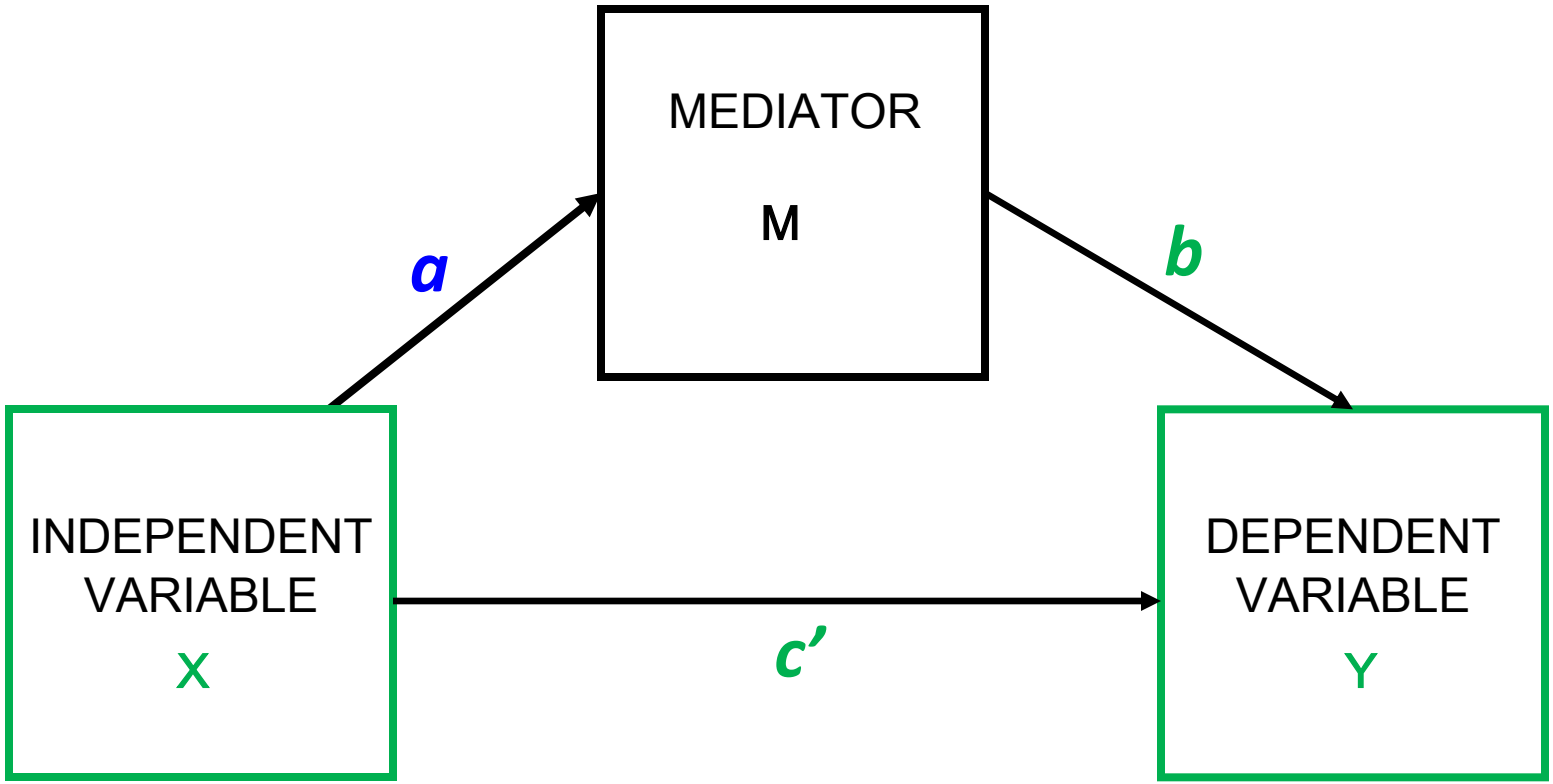
Mediator

A variable that is intermediate in the causal process relating an independent to a dependent variable.

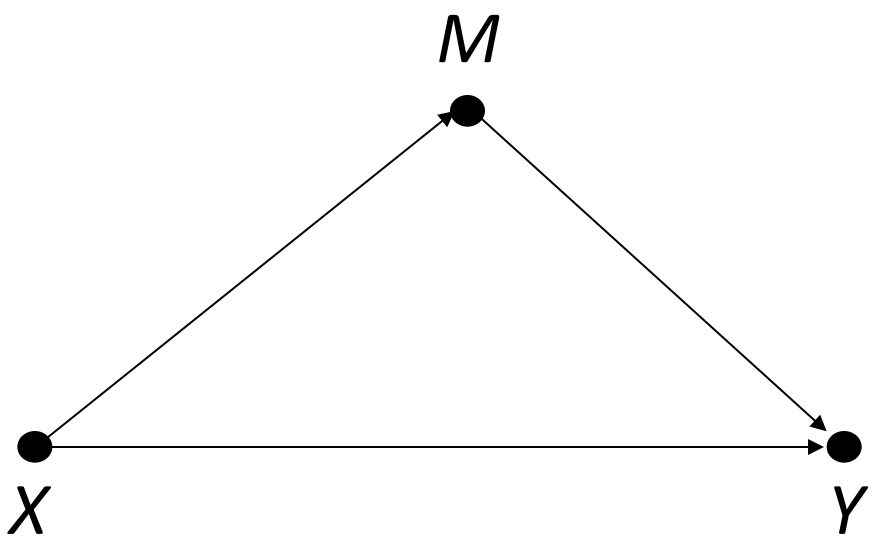
<u>Some Examples:</u>

1. Neglect/Abuse in childhood (X) to **impaired threat appraisal** (M) to aggressive behavior in adolescence (Y) Dodge, Bates, & Pettit (1990).
2. Tobacco prevention program promotes **anti-tobacco norms** which reduce tobacco use (MacKinnon et al., 1991)
3. Screening program increases **identification of early stage cancer** which reduces cancer deaths (Zauber, 2015).
4. Adverse Childhood Experiences affect **perinatal risk** which affects maternal-infant dyadic functioning (Roubinov et al., 2021).
5. Your Examples?

Single Mediator Model



Mediation Directed Acyclic Graph (DAG)



Mediator Definitions

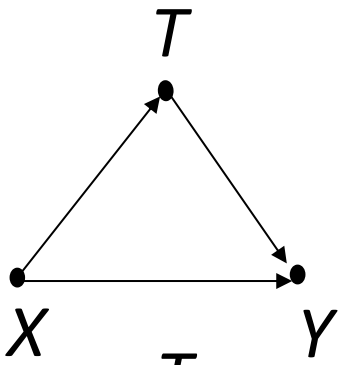
- A mediator is a variable in a chain whereby an independent variable causes the mediator which in turn causes the outcome variable (Sobel, 1990).
- A variable that occurs in a causal pathway from an independent variable to a dependent variable. It causes variation in the dependent variable and itself is caused to vary by the independent variable (Last, 1988).

Two, Three, Four variable effects

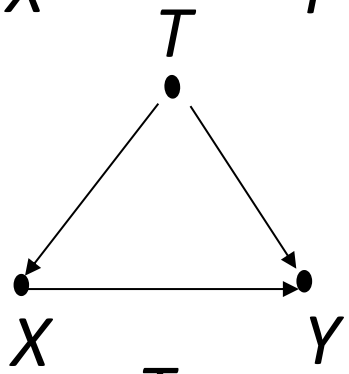
- Two variables: $X \rightarrow Y$, $Y \rightarrow X$, $X \leftrightarrow Y$ are reciprocally related. Measures of effect include the correlation, covariance, regression coefficient, odds ratio, mean difference.
- Three variables: $X \rightarrow M \rightarrow Y$, $X \rightarrow Y \rightarrow M$, $Y \rightarrow X \rightarrow M$, and all combinations of reciprocal relations. Special names for third-variable effects: confounder, mediator, collider.
- Four variables: many possible relations among variables, e.g., $X \rightarrow Z \rightarrow M \rightarrow Y$

Third-Variable (T) Effects

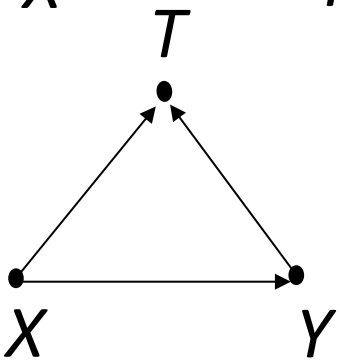
Mediator: This is the focus of this talk. A variable that is intermediate in a causal process between X and Y.



Confounder: A variable that causes X and Y such that if it is not included in the analysis an incorrect estimate of the relation between X and Y will be obtained.



Collider: A variable that is caused by X and Y so that it should not be adjusted in the analysis of X and Y because it will incorrectly change the relation between X and Y.



Mediation is important because …

- Theoretical and applied questions in many fields are about mediating processes
- Extracts more information from a research study
- Critical for applied research, especially prevention and treatment to identify critical ingredients leading to more efficient interventions

New Focus on Mediating Mechanisms

"First, future trials will follow an experimental medicine approach in which interventions serve not only as potential treatments, but as **probes to generate information about the mechanisms underlying a disorder…** It offers us a way to understand the **mechanisms by which these treatments are leading to clinical change**."

Thomas Insel, M.D. NIMH Director:
http://www.nimh.nih.gov/about/director/2014/a-new-approach-to-clinical-trials.shtml

Developmental Mediation Questions

- What are the mechanisms by which education, parental involvement, maltreatment, and environment during childhood affect educational achievement, and physical and mental health in later life (e.g., Braveman & Barclay, 2009)?
- How does poverty/neglect/abuse in childhood affect behavior in adolescence and adulthood?
- How can predictors of developmental outcomes be used to design and refine interventions to reduce problem behavior?

Applications of Mediation

Two overlapping applications of mediation analysis:

1. **Mediation for Explanation**

2. **Mediation by Design**

Mediation for Explanation

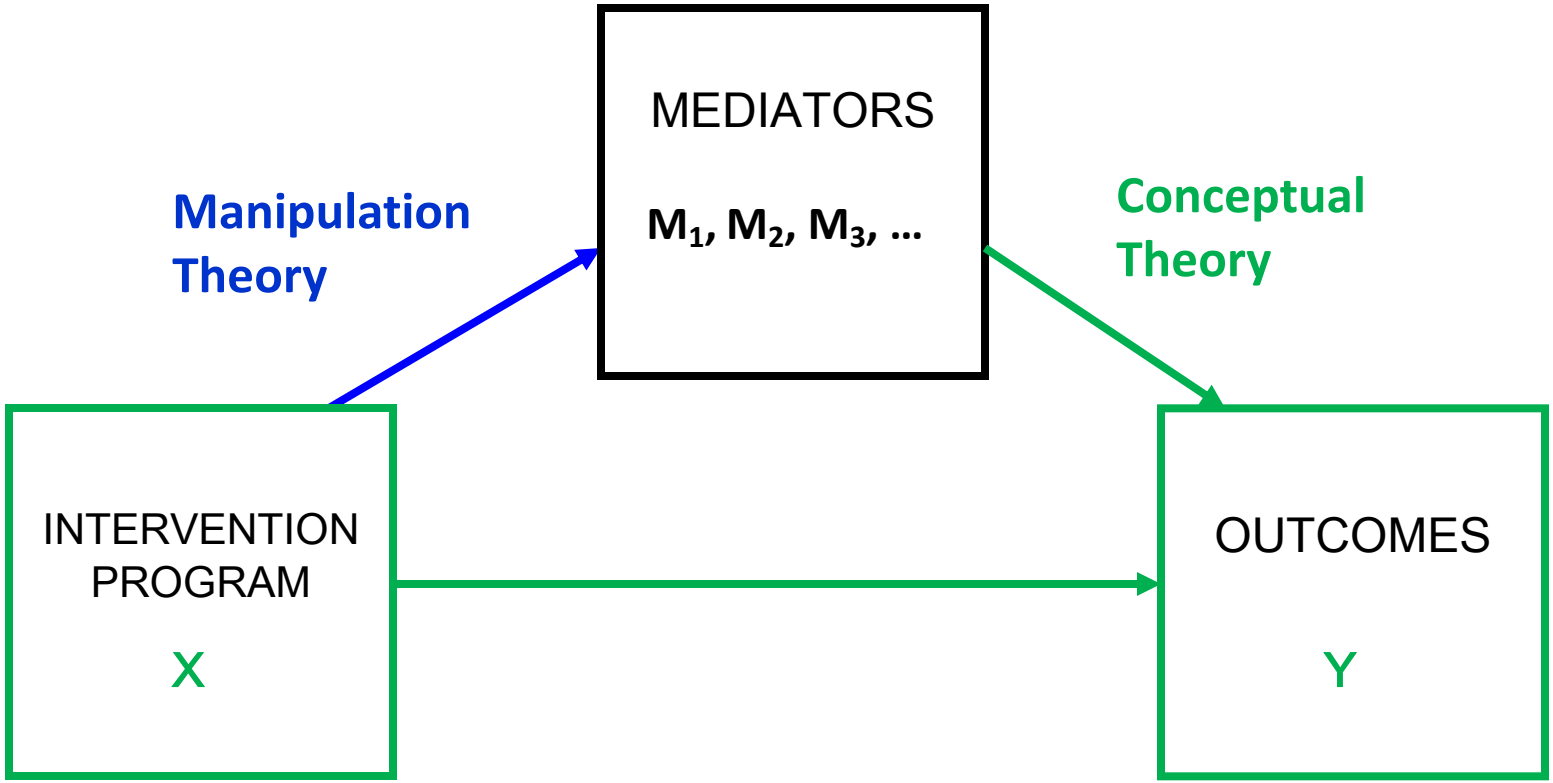
- Observe a relation and then try to explain it with a mediating process.

- Elaboration method described by Lazarsfeld and colleagues (1955; Hyman, 1955) where third variables are included in an analysis to see if/how the observed relation changes.

Mediation by Design

- Select mediating variables that are causally related to an outcome variable.
- Intervention is designed to change these mediators.
- If mediators are causally related to the outcome, then an intervention that changes the mediator will change the outcome.
- Common in applied research like prevention and treatment.

Intervention Mediation Model



If the mediator changed is causally related to Y, then changing the mediator will change Y.

Mediation in Intervention Research

- A theory based approach focuses on the processes underlying programs. Mediators play a primary role. **Manipulation Theory** corresponds to how the manipulation will affect mediators. **Conceptual Theory** focuses on how the mediators are related to the dependent variables (Chen, 1990; Lipsey, 1993; MacKinnon, 2008).
- Identifying mediators is important for basic and applied science. Practical implications include reduced cost and more effective interventions if true mediators are identified.

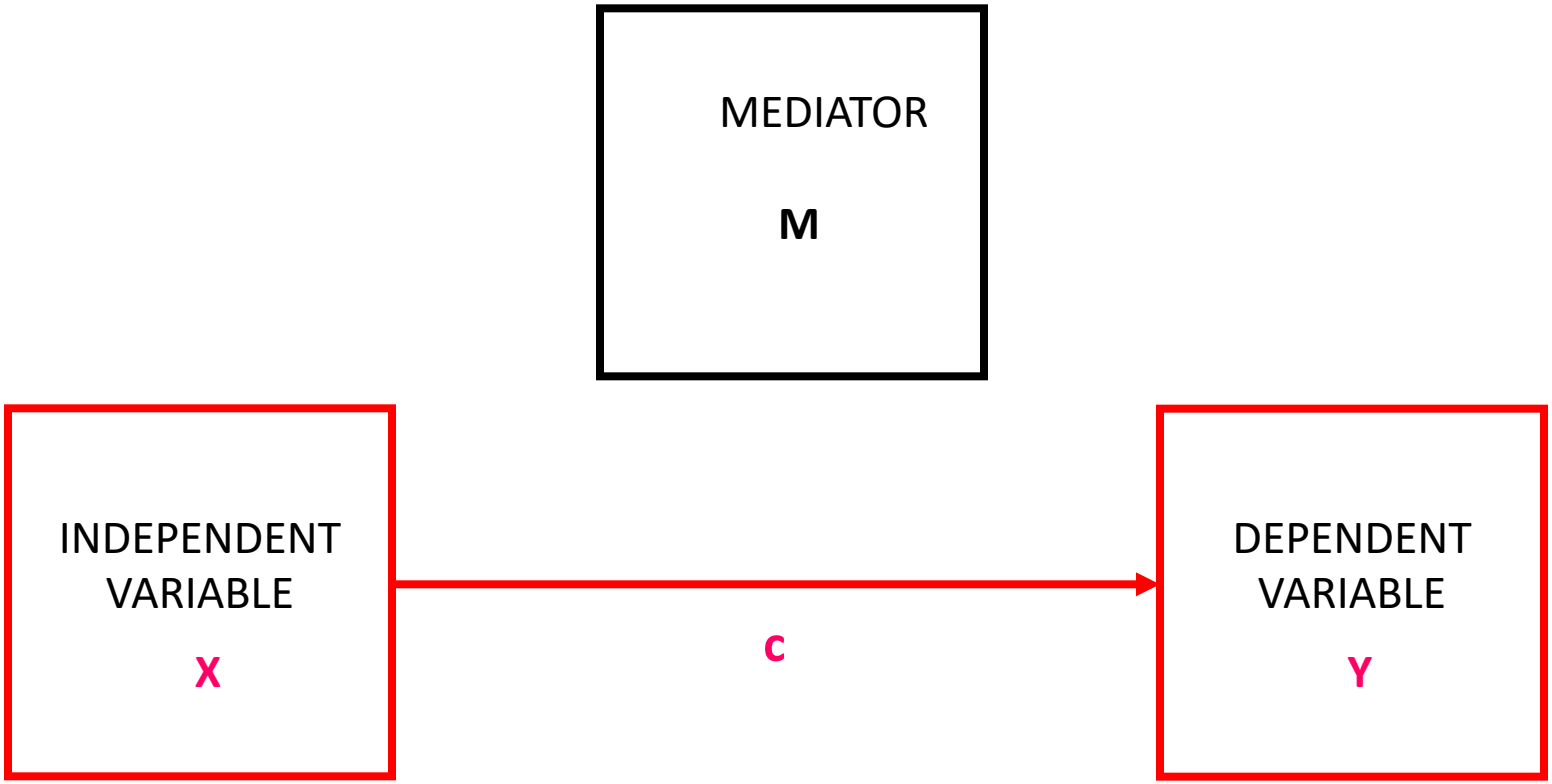
Interventions and Mediators

- Home visitation intervention (X) focusing on life course variables such as the number of subsequent children born to mothers (M1) and use of public assistance (M2) led to reductions in child maltreatment (Y) (Eckenrode et al., 2016, *Child Maltreatment*).
- Head Start (X) targets enhancing executive function (M) to improve cognitive and emotional skills (Y) (Bierman et al., 2008, *Development and Psychopathology*)
- Early Head Start (X) affects child's engagement in activities (M1) and child's cognitive development (M2) which reduced Maltreatment (Y) (Green et al., 2018, CDC Report).

Mediation Regression Equations

- Tests of mediation for a single mediator use information from some or all of three equations.
- The coefficients in the equations may be obtained using methods such as ordinary least squares regression, covariance structure analysis, or logistic regression.
- The product of coefficients test is the method of choice. It extends to more complicated models such as the multiple mediator model.

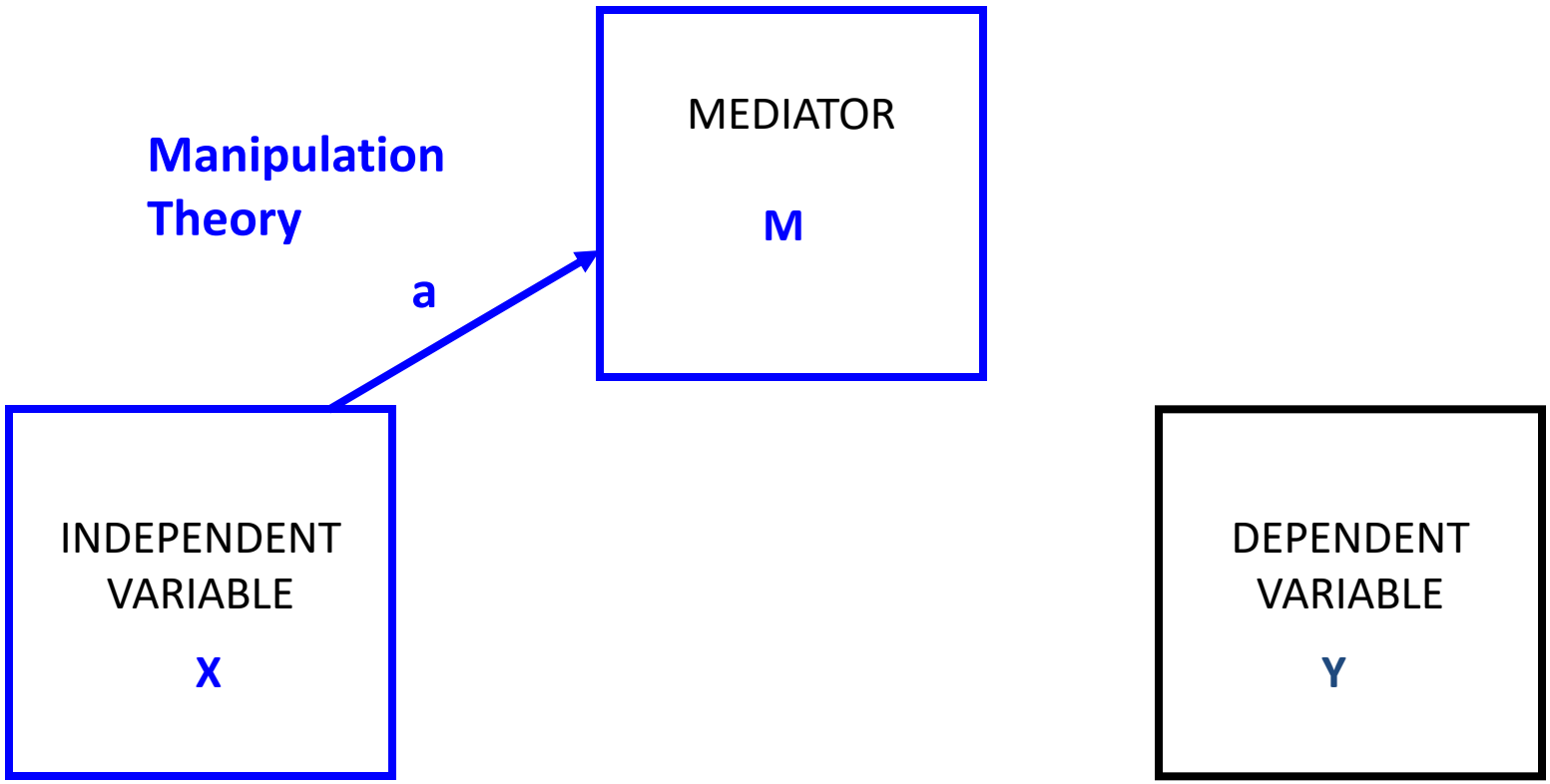
Regression Equation 1



1. The independent variable is related to the dependent variable:

$$Y = i_1 + \hat{c}X + e_1$$

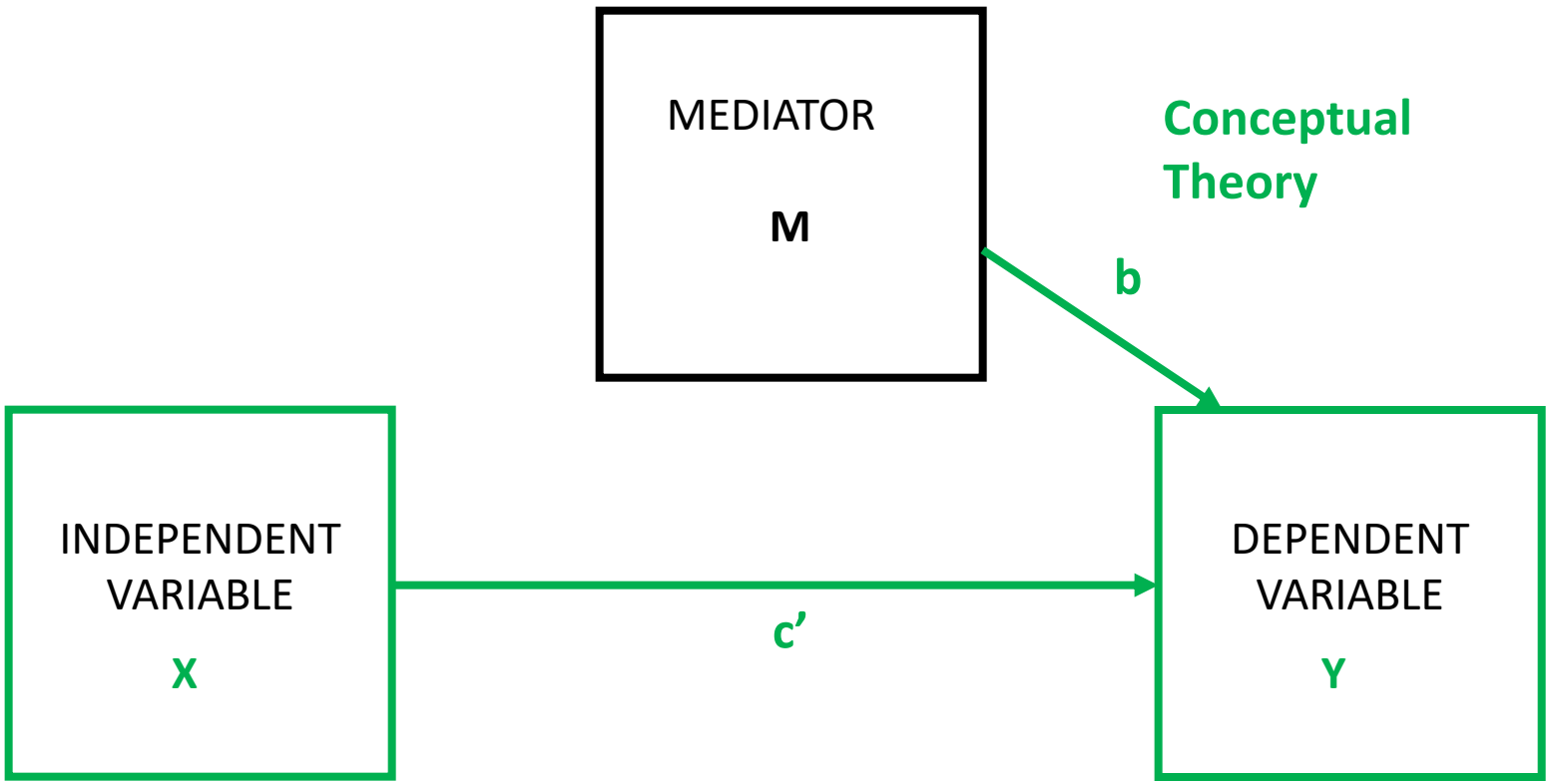
Regression Equation 2



2. The independent variable is related to the potential mediator:

$$M = i_2 + \hat{a}X + e_2$$

Regression Equation 3



3. The mediator is related to the dependent variable controlling for exposure to the independent variable:

$$Y = i_3 + \hat{c}' X + \hat{b} M + e_3$$

Mediated Effect Measures

Mediated effect = ab Product of Coefficients

Mediated effect = c-c' Difference in Coefficients

Mediated effect = ab = c-c'

(see MacKinnon et al., 1995 for a proof)

Direct effect = c' & Total effect = ab+c' = c

Mediated Effect, $\hat{a}\hat{b}$, Standard Error

Mediated effect = $\hat{a}\hat{b}$, Standard error = $\sqrt{\hat{a}^2 s_{\hat{b}}^2 + \hat{b}^2 s_{\hat{a}}^2}$

Multivariate delta method standard error
(Sobel 1982)

Test for significant mediation:

$$z' = \frac{\hat{a}\hat{b}}{\sqrt{\hat{a}^2 s_{\hat{b}}^2 + \hat{b}^2 s_{\hat{a}}^2}}$$

Compare to empirical distribution of the mediated effect

Assumptions

- Reliable and valid measures
- Coefficients, *a*, *b*, *c'* reflect true causal relations and the correct functional form. No omitted influences.
- Mediation chain is correct: Temporal ordering is correct. X before M before Y.
- Homogeneous effects across subgroups: The relation from X to M and from M to Y are homogeneous across subgroups of participants in the study, i.e., No moderators.

For each method of estimating the mediated effect based on Equations 1 and 3 (*c* - *c'*) or Equations 2 and 3 (*ab*):

Significance Testing and Confidence Limit Estimation

- Product of coefficients estimation of the mediated effect, *ab*, and standard error is the most general approach with best statistical properties.
- Best tests are the **Joint Significance**, **Distribution of the Product**, and **Bootstrap** for confidence limit estimation and significance testing (MacKinnon et al., 2004; 2007).

Empirical Sample size estimates for .8 power to detect the mediated effect

Test	S-S	S-M	M-M	L-L
Causal Steps ($c' = 0$)	**20886**	3039	397	92
Normal	667	422	90	42
Dist. Product	539	401	74	35
Bootstrap	558	406	78	36

Note: N required for a complete mediation model, $c' = 0$. Table entries are based on empirical simulation so they are not exact (Fritz & MacKinnon, 2007). S=small (r~=.10, M= medium (r~=.3), and L=large (r~=.5) approximate effect size. For example, S-S means small effect size for the *a* path and small effect size for the *b* path.

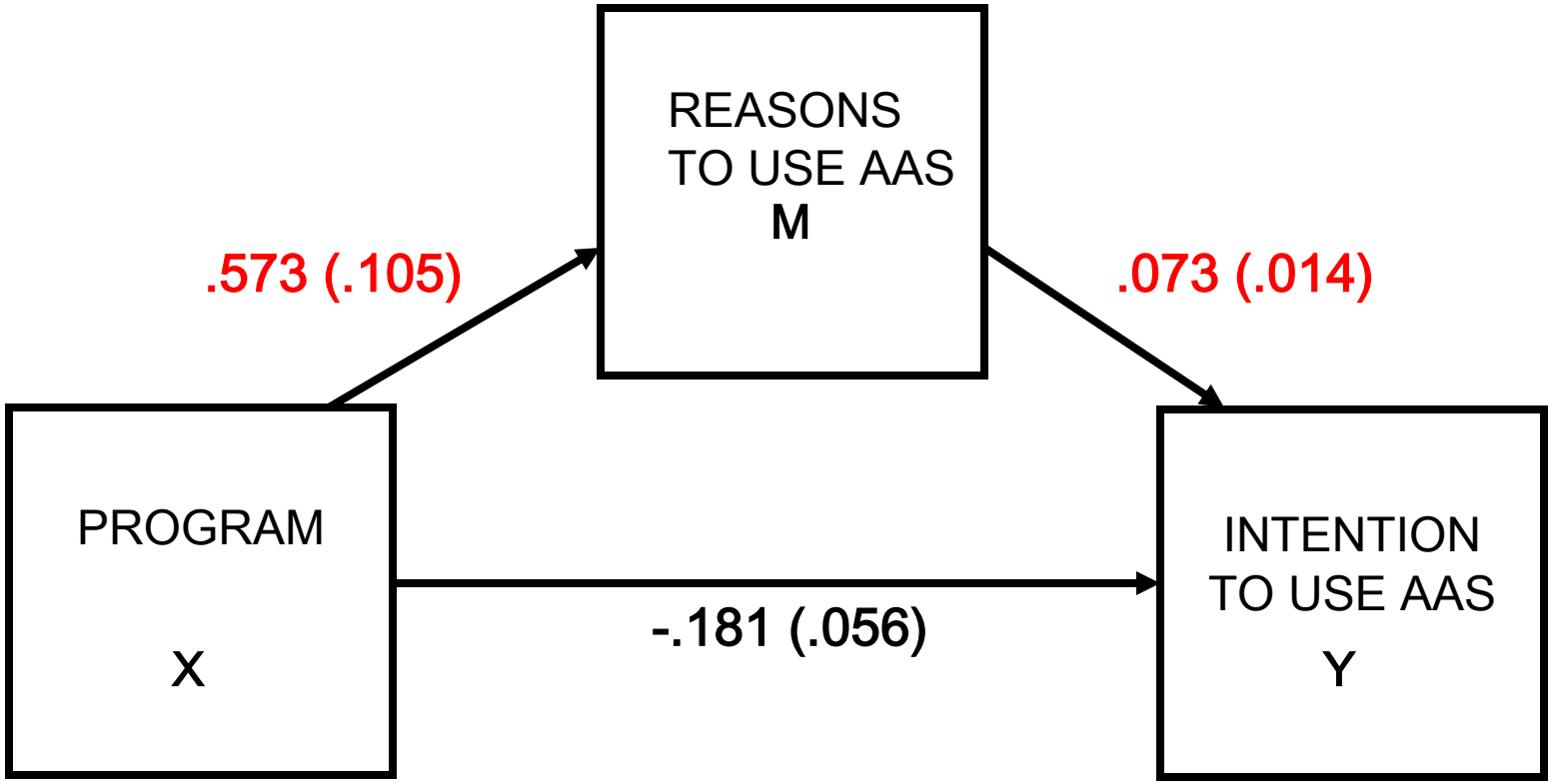
Mediation and Nonsignificant X on Y Effect

- It is important to conduct mediation analysis whether an overall effect of X on Y is statistically significant or not.
- It is possible to obtain a nonsignificant overall effect of X on Y when there is statistically significant mediation (O'Rourke & MacKinnon, 2015).
- Mediation analysis provides information about **Manipulation theory** (X on M) and **Conceptual Theory** (M on Y). Failure of one or both theories could lead to a nonsignificant effect of X on Y.

Inconsistent Mediation Models

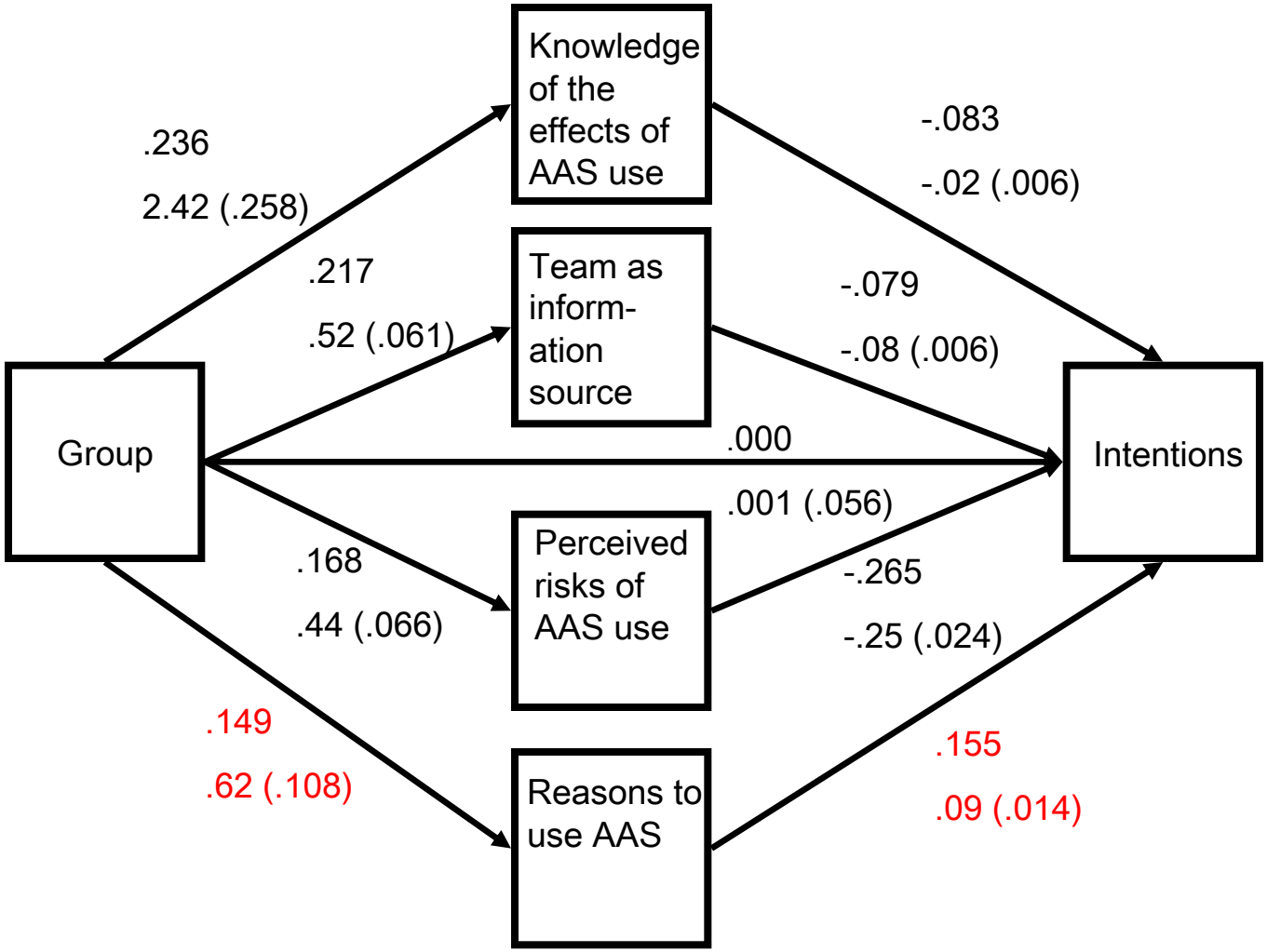
- An inconsistent mediation model has at least one mediated effect with a different sign than the direct effect or other mediated effects (MacKinnon et al., 2000)
- There is mediation because the mediator transmits the effect of the independent variable to the dependent variable.
- An inconsistent mediation model can occur whether or not $\hat{c}$ is statistically significant.
- Intervention studies may have a mediator that is counterproductive. The best way to find these variables is to use mediation analysis.

Inconsistent Mediation in a Steroid Prevention Study

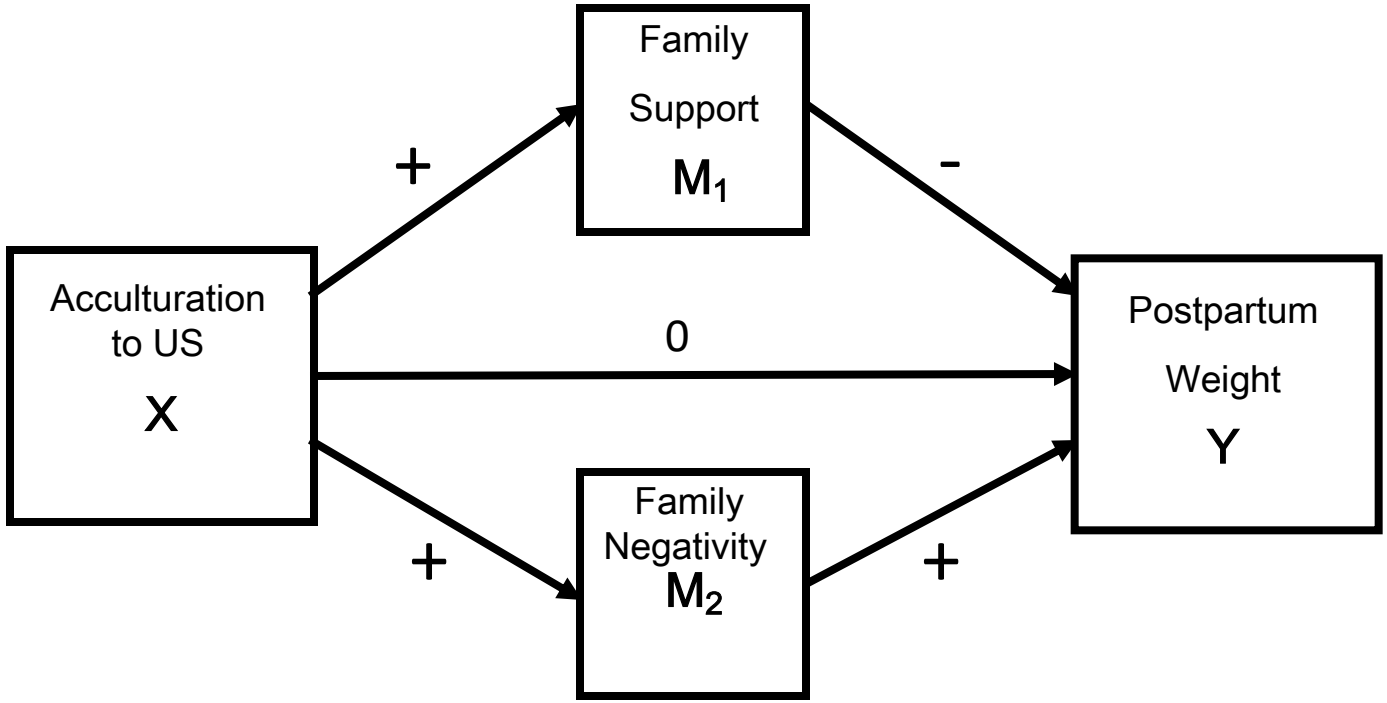


Mediated effect = .042
Standard error = .011

Multiple Mediator Model of Intent to Use Anabolic Steroids



Mediators of Acculturation on Weight (Jewell et al. , 1984)



Acculturation increases Family Support and Family Negativity but Family Support reduces Weight and Family Negativity increases Weight.

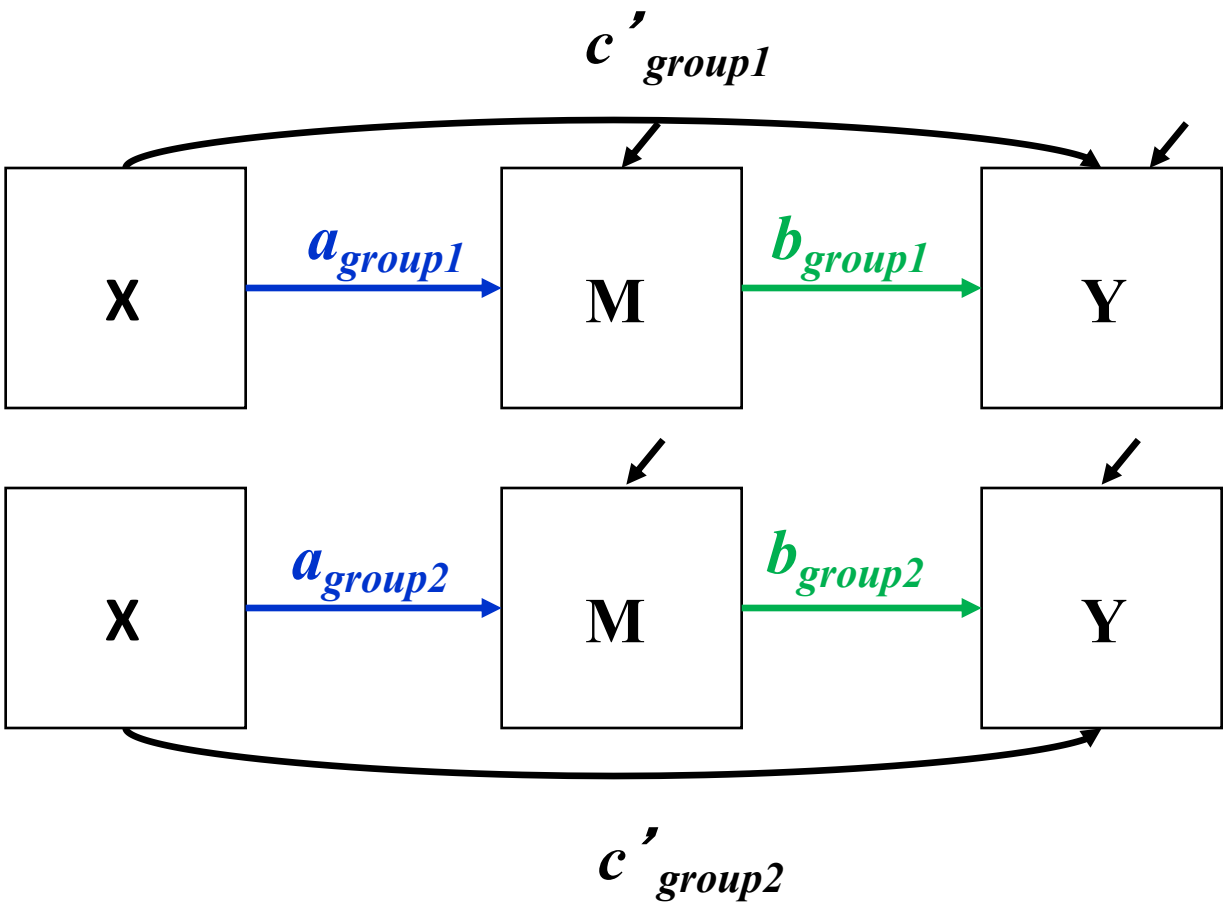
Path Model for Testing Homogeneity of Effects (Moderation) across Groups

GROUP 1

Mediated effect:
$a_{grp1}b_{grp1}$

GROUP 2

Mediated effect:
$a_{grp2}b_{grp2}$



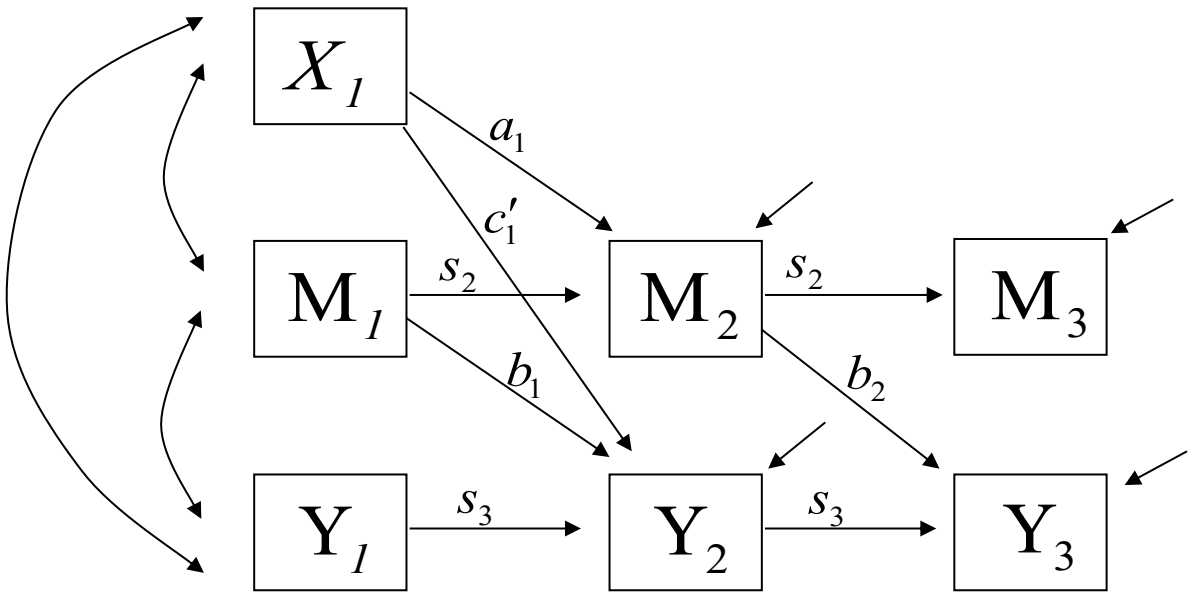
Longitudinal Mediation Analysis

- Mediation is a longitudinal model. Assume correct temporal ordering: X before M before Y.
- Relations among X, M, and Y are at some equilibrium so the observed relations are not solely due to when they are measured, i.e., if measured 1 hour later a different model would apply.
- Assumes correct timing and spacing of measures to detect effects.
  - When does X affect M and M affect Y
  - Triggering, cascading, and other timing processes may be at work (Tang & DeRubeis, 1999; Howe et al., 2002)
  - Timing is crucial for deciding when to collect longitudinal measures (Collins & Graham, 2002)

What if Repeated Measures of X, M, and Y are Available?

- Measures of X, M, and Y at two time points: Difference Score, ANCOVA, Residualized Change Score.
- Measures of X, M, and Y at three or more time points: Autoregressive, Latent Growth, Latent Change Score Models, Survival Models, and methods to reduce to a few measures, e.g. Area Under the Curve.
- For intervention research, X is usually measured once and represents random assignment of participants to one of two groups.

Autoregressive Model with Time-Ordered Mediation



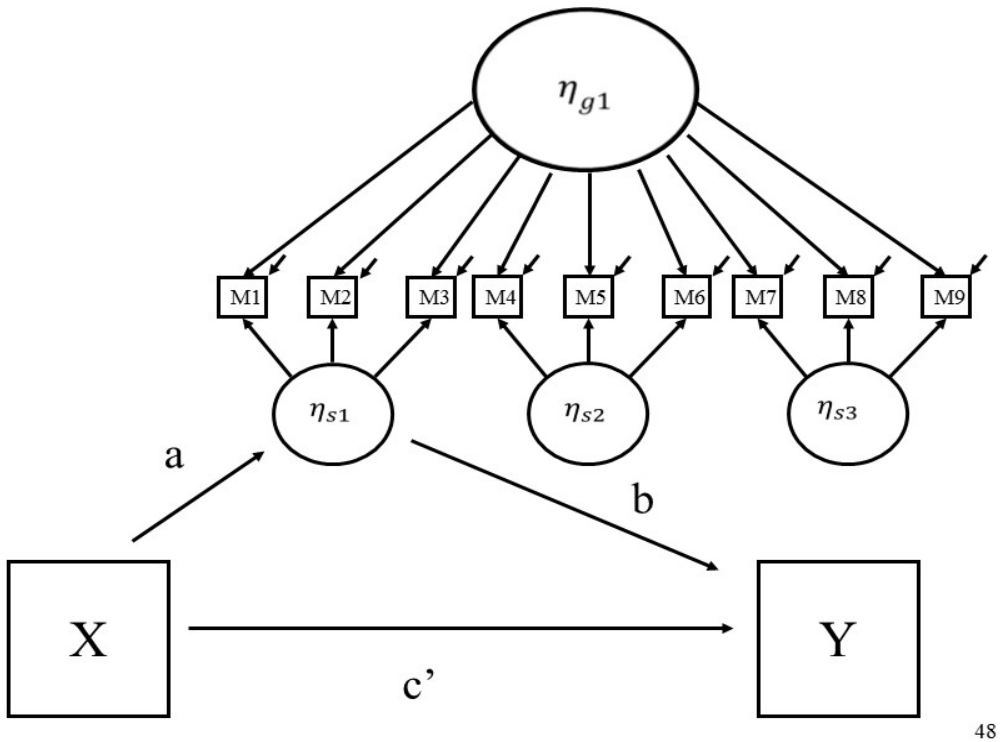
Note: All residuals are correlated

(Cole & Maxwell, 2003; MacKinnon, 1994; 2008)

Mediation as a Measurement Challenge

- Progress is made with more accurate measurement of the mediating process.
- Genetic Mediation Theory (Parent $\rightarrow$? $\rightarrow$ Offspring), Atomic Mediation (Input $\rightarrow$? $\rightarrow$ Output) Theory, Causes of Disease
- Refinement of measurement to identify the most critical facet of constructs. The facet that is the critical ingredient.
- Many mediating constructs are multidimensional. The task is to evaluate intervention effects through facets.

Bifactor Mediation Model



- Facet of a construct is the important mediator in this model (Gonzalez & Mackinnon, 2017)

Causal Inference in Mediation

- Methods described previously assume true causal relations and no omitted variables for mediation analysis.

- Blalock's (1979) presidential address– about 50 variables are involved in sociological phenomenon. Comprehensive health psychology models. How many variables are relevant for child maltreatment research?

- Problem with mediation analysis because M is not randomly assigned but is self-selected.

Counterfactual/ Potential Outcome Models

- Most modern causal inference approaches are based on a counterfactual or potential outcomes framework.
- In these models, all the possible counterfactual and actual conditions of an experiment are considered and the statistical model is based on all these possible or potential conditions.
- Requires consideration of conditions that did not occur.
- Counterfactual thinking is common, e.g., If I had three donuts instead of none this morning, I would have more energy.

Observed and Counterfactual Table

Participant	Assignment	Observed Y	Potential Outcome Y(0)	Potential Outcome Y(1)
Juan	0	0	0	?
Julie	0	1	1	?
Alondra	0	1	1	?
Antonio	0	1	1	?
Kim	1	1	?	1
Derrick	1	0	?	0
Reginald	1	0	?	0
Shar	1	0	?	0

Randomized Two Group Design

- Ideally we need the same individual in both the treatment and control conditions at the same time. Units (individual level) usually have observed data for one of two conditions but not the other—the fundamental problem of causal inference (Holland, 1986).
- Randomization of a large number of persons resolves the fundamental problem of causal inference. The average in each group can be compared and is an estimator of a causal effect. It is called an average causal effect (ACE).

Counterfactual/ Potential Outcome Models

- For the two group study, you have the data for a person who serves in the control group $\mathbf{Y_i(0)}$ but you do not have the data for the same person in the treatment group $Y_i(1)$. Similarly, you have data for a person in the treatment group $\mathbf{Y_i(1)}$ but not for that person in the control group $Y_i(0)$.
- Solution is to take averages in each group after randomization Causal Effect = E(Y(0))-E(Y(1))
- For mediation, M must also be considered: E(Y(0,M(0)), E(Y(0,M(0)), E(Y(1,M(0)), E(Y(0,M(1)).
- Overall goal is to use the information that you do have, assumptions, observed data, to gain insight all potential outcomes.

Causal Effects Compare Potential Outcomes

Total Natural Indirect Effect

TNIE = E[Y(1,M(1)) – Y(1,M(0))] = *ab* + *ah*

Pure Natural Indirect Effect

PNIE = E[Y(0,M(1)) – Y(0,M(0))] = *ab*

Software is available to estimate causal mediation effects in Mplus, SAS CAUSALMed, R (Mediation and Medflex), Stata (Paramed)

Note: *h* is the interaction effect of X and M on Y. See MacKinnon et al., (2020, *Prevention Science*) for links between traditional and causal mediation.

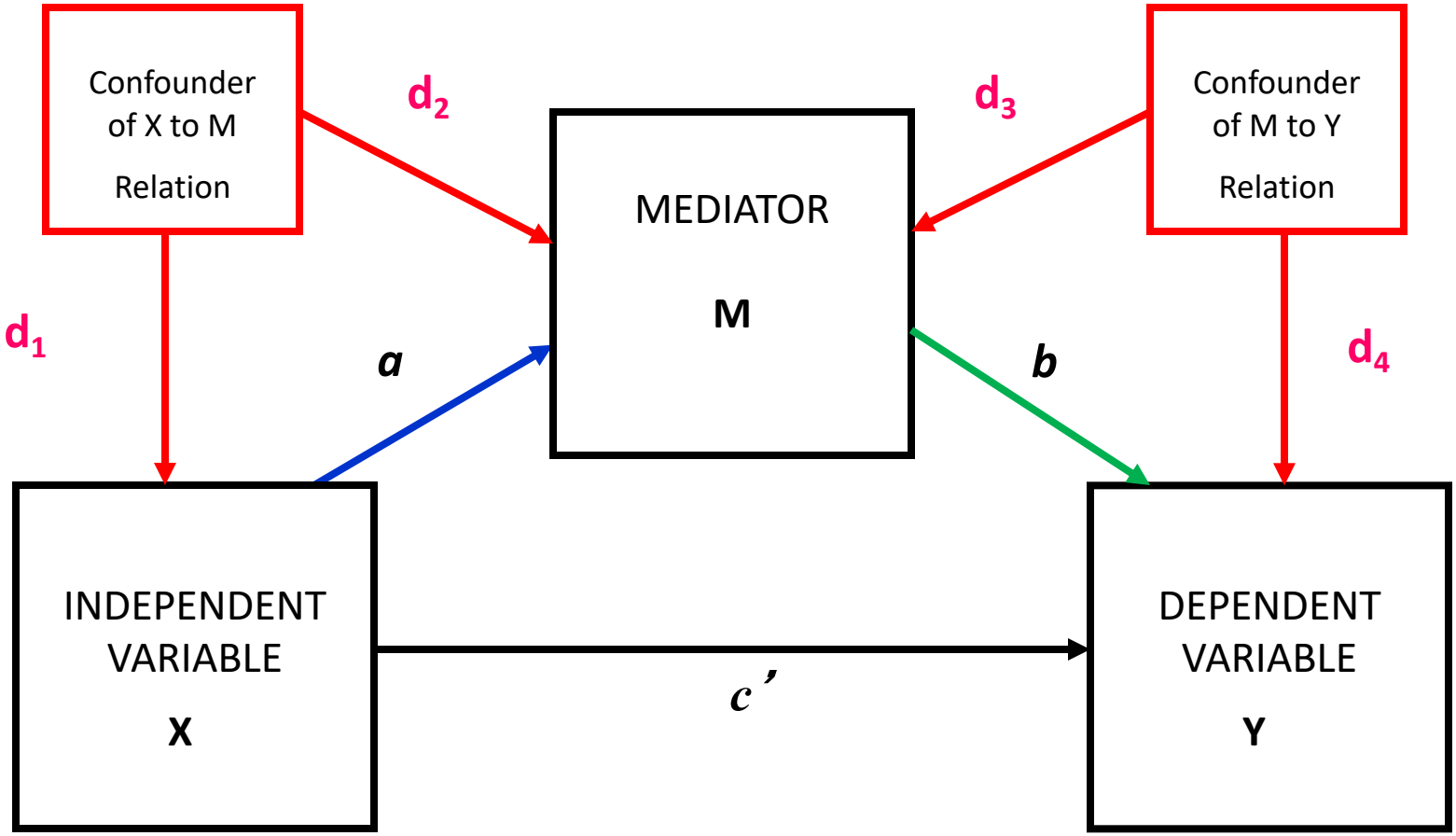
Confounding Assumptions

1. No unmeasured X to Y confounders given covariates.
2. No unmeasured M to Y confounders given covariates.
3. No unmeasured X to M confounders given covariates.
4. There is no effect of X that confounds the M to Y relation.

(VanderWeele & Vansteelandt, 2009)

- Randomized X satisfies Assumptions 1 and 3 but not 2 and 4.

Confounders of Mediation Relations



True model needs d_1,d_2, d_3, d_4, otherwise coefficients are confounded.

45

Sensitivity Analysis for Confounding

How will results change with confounding of the M to Y relation, e.g. when X is randomized?

- Adaptation of Left Out Variables Error (LOVE; Mauro, 1990) based on the correlation of a confounder with Y and the correlation of a confounder with M.
- VanderWeele (2010), confounder effect on Y and difference in proportions of the confounder between groups at level of M, called E-value.
- Imai et al. (2010), confounder effect as the correlation between error terms.
- See Cox et al., 2014, *Evaluation Review*.

Statistical Methods for Confounding

- Statistical approaches to improve causal inference from a mediation study. A way to deal with confounder bias.

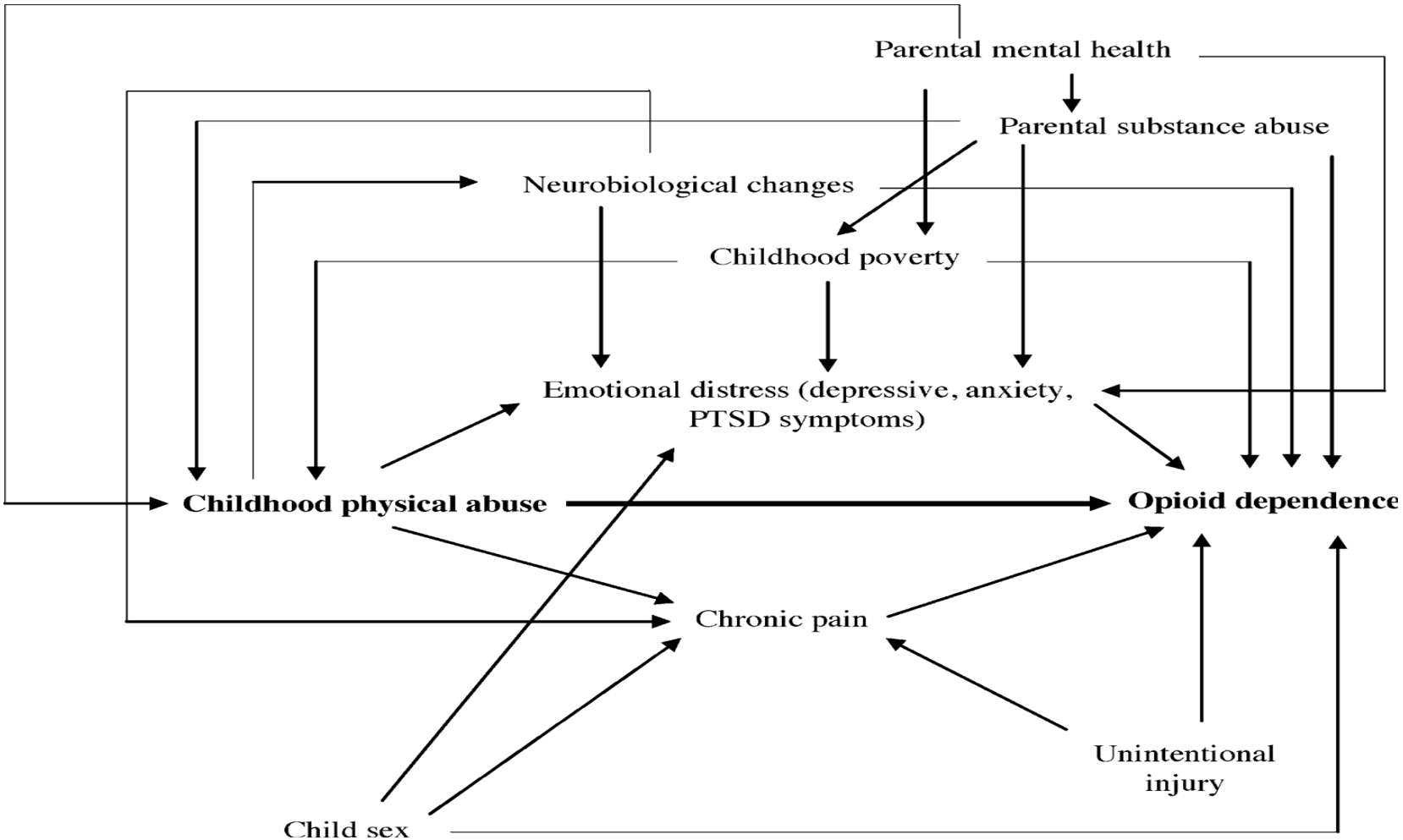
  1. Instrumental Variable Methods
  2. Principal Stratification
  3. Inverse Probability Weighting
  4. G-estimation

- Active area of research (MacKinnon & Pirlott, 2015, *Personality and Social Psychology Review*)…

Structural Causal Model (SCM)

- General method to specify causal relations. Similar to Structural Equation Modeling but based on graph theory and Bayesian networks from computer science.
- Probabilistic graph models that represent dependencies among random variables. Used by computers to compute conditional probabilities.
- Extended to causal inference by Judea Pearl.
- Widely used in epidemiology at least part because of links with work by Jamie Robins, Sander Greenland, and Donald Rubin.

Austin et al. (2019, p. 82, Figure 5) DAG. *Child Abuse and Neglect*



Application of SCM to Childhood Poverty and Adult Income

- Bellani & Bia (2018) applied modern causal inference methods to estimate the mediating effect of education on the relation of childhood poverty to economic outcomes. Applied causal mediation estimators and evaluated sensitivity of results to unmeasured confounders.
- Poverty in childhood affects secondary education which accounted for about 30% of the total effect on adult income.

Bellani & Bia (2018). *Journal of the Royal Statistical Society*.

Design Approaches to Improving Causal Inference

- What is the best next study or studies to conduct after a statistical mediation analysis to test mediation theory.

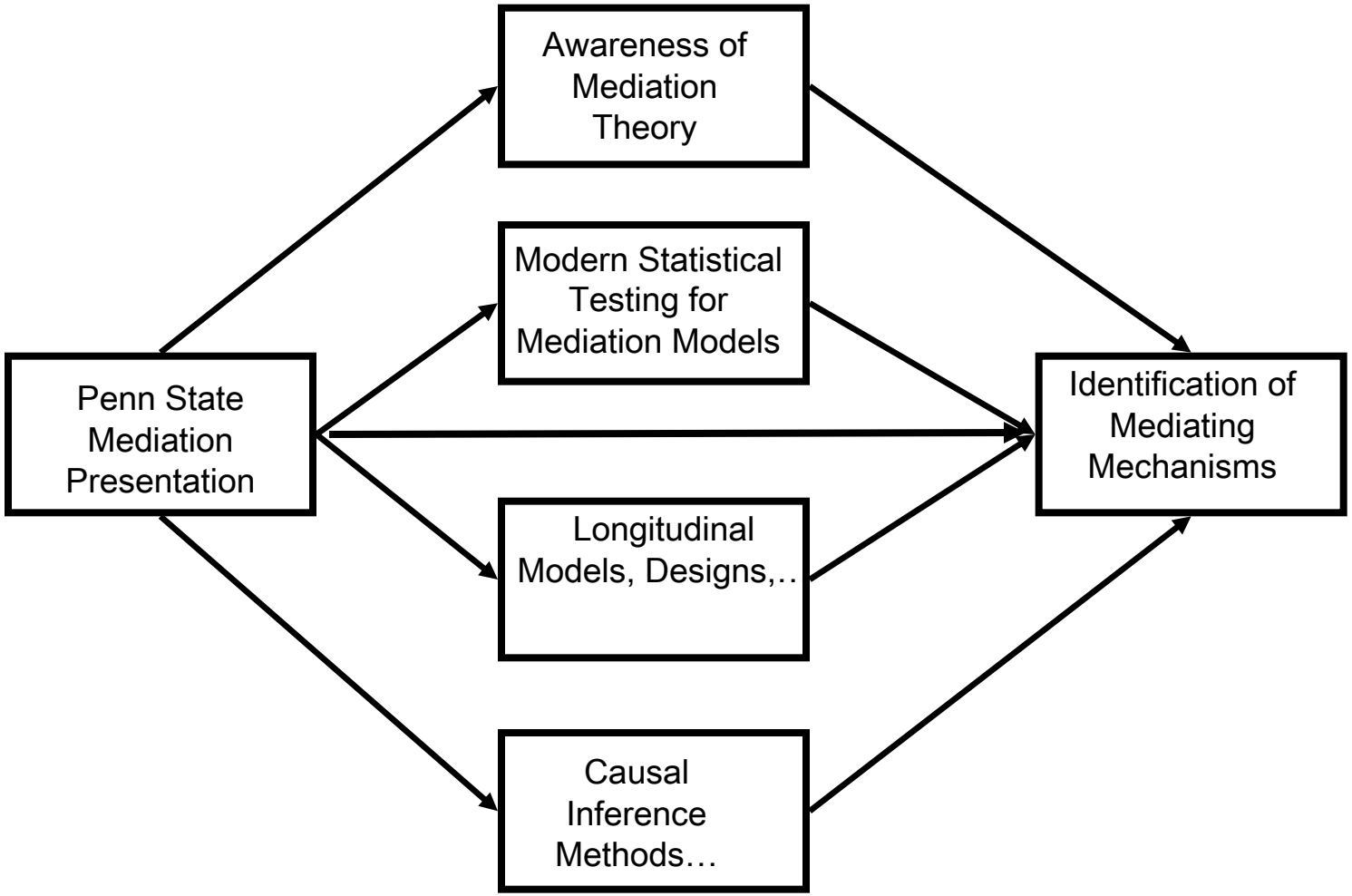
1. Designs to address **Consistency** of the mediation relation.
2. Designs to address **Specificity** of the mediation relation.

MacKinnon, 2008; MacKinnon, Cheong, & Pirlott, 2013 related to Hill's (1971) considerations. Pirlott & MacKinnon, 2017; Also SMART designs (Almirall et al., 2014)

Summary

- Consider mediation (and confounder and collider) effects, in addition to two variable effects.
- Mediation is important for understanding how effects come about for experimental and observational studies.
- There are challenging assumptions for accurate conclusions, especially related to making causal conclusions.
- Multiple Mediator Models, Models with Moderation and Mediation, Longitudinal Mediation Models and other models are available.
- Causal inference is an active research area generating new methods to investigate confounder bias and other threats to accurate conclusions.

Hypothesized Effects of Penn State Mediation Analysis Presentation



Thank You

References available by contacting David.MacKinnon@asu.edu

Collaborators

Leona Aiken, Amanda Baraldi, Hendricks Brown, Jeewon Cheong, Donna Coffman, Matt Cox, Stefany Coxe, James Dwyer, Diane Elliot, Craig Enders, Amanda Fairchild, Matt Fritz, Lois Gelfand, Linn Goldberg, Kim Goldsmith, Oscar Gonzalez, Kevin Grimm, Jeanne Hoffman, Booil Jo, David Kenny, Yasemin Kisbu-Sakarya, Jennifer Krull, Kerry Kuehl, Linda Luecken, Ginger Lockhart, Chondra Lockwood, Maria Maric, Gina Mazza, Milica Miocevic, Antonio Morgan-Lopez, Vanessa Ohlrich, Magarita Olivera-Aguilar, Holly O'Rourke, Will Pelham, Mary Ann Pentz, Angela Pirlott, Krista Ranby, Judith Rijnhart, Mark Reiser, Heather Smyth, Elizabeth Stuart, Marcia Taborga, Aaron Taylor, Jenn Tein, Felix Thoemmes, Davood Tofighi, Matt Valente, Liz Stuart, Martia van Stralen, Wei Wang, Ghulam Warsi, Steve West, Jason Williams, Ingrid Wurpts, Mine Yildrim, Myeongsun Yoon, Ying Yuan et al.

Recommendations

- Mediation is important for understanding how effects come about for experimental and observational studies.
- There are challenging and controversial assumptions for accurate conclusions.
- Multiple Mediator Models, Models with Moderation and Mediation, Longitudinal Mediation Models and other models are available.
- Causal inference is an active research area generating new methods to investigate confounder bias and other threats to accurate conclusions Focus on the mediating process by which one variable is related to another variable. Specify theory for mediating variables
- Consider the difference between short term and long term mediation.
- Consider mediation (and confounder and collider) effects, in addition to two variable effects.
- When designing an intervention consider how difficult it will be to change the mediator.