

# Mediating Mechanisms in Intervention Research

David P. MacKinnon, PhD  
Arizona State University

This research was supported by the National Institute on Drug Abuse and the  
National Institute on Mental Health

Presented at the 2009 CHIPS Summer Research Institute in Tempe, AZ



# Outline of the Talk and Mediation Resources

- Talk Overview:
  1. Overview of Mediation Analysis
  2. Statistical Tests for Mediation
  3. Longitudinal Mediation Models
  4. Designs to Assess Mediation
- Website: [www.public.asu.edu/~davidpm/](http://www.public.asu.edu/~davidpm/)
- New Book: MacKinnon, D. P. (2008). Introduction to Statistical Mediation Analysis. Mahwah, NJ: Erlbaum.
  - <http://www.psypress.com/9780805839746>



# Mediator

- A variable that is intermediate in the causal process relating an independent to a dependent variable:
  - Program has beneficial effects on cholesterol, smoking, and high blood pressure to reduce cardiovascular disease (MRFIT Group, 1990).
  - Tobacco prevention program promotes anti-tobacco norms which reduces tobacco use (MacKinnon et al., 1991).
  - Screening program increases identification of early stage cancer which reduces cancer deaths.
  - Wellbutrin (Bupropion) increases subject's willingness to quit and self-efficacy which are associated with one month abstinence from tobacco (McCarthy et al., 2008).



# Mediator Definitions

- A variable in a chain whereby an independent variable causes the mediator which in turn causes the outcome variable (Sobel, 1990).
- The generative mechanism through which the focal independent variable is able to influence the dependent variable (Baron & Kenny, 1986).
- 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).



# More Mediation Examples

- Psychotherapy induces catharsis, insight, and other mediators which lead to a better outcome (Freedheim & Russ, 1981).
- Psychotherapy changes attributional style which reduces depression (Hollon, Evans, & DeRubies, 1991).
- Parenting programs reduce parents' negative discipline which reduces symptoms among children with ADHD (Hinshaw, 2002).

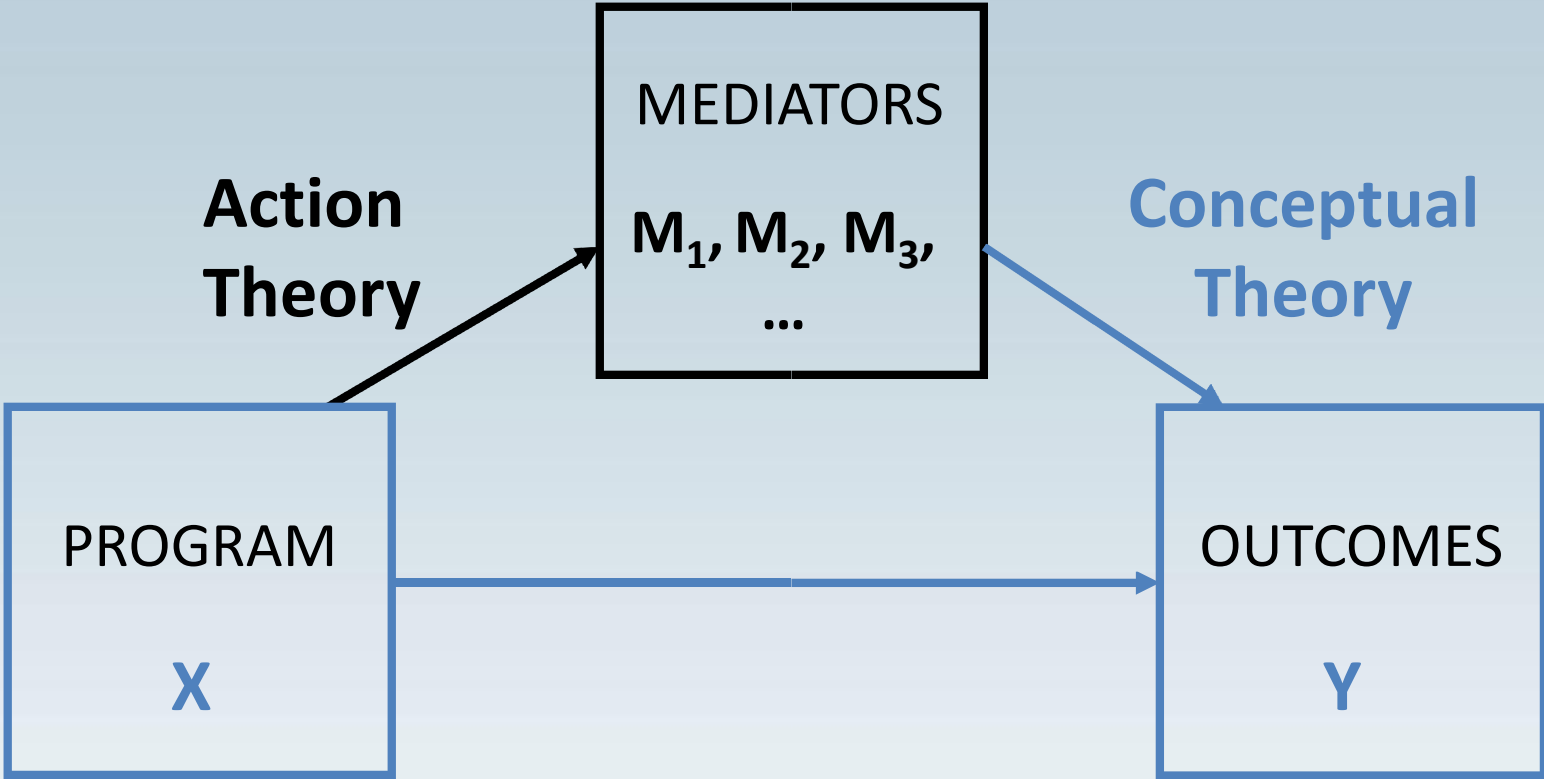


# General Intervention Mediation Model

- There is some form of an intervention, often a randomized intervention, where some participants receive the intervention and others do not. The intervention itself is often composed of many components or sessions.
- There are measures of major outcomes. These may be continuous (depression, blood pressure, achievement) or categorical (drug use, death, screening).
- There is often theory for how the intervention leads to beneficial effects on the outcome measures through mediating mechanisms.



# Intervention Mediation Model



# Quotes

- In the absence of a concern for such mediating or intervening mechanisms, one ends up with facts, but with incomplete understanding (Rosenberg, 1968, p. 63).
- Nursing “.. Should consider hypotheses about mediators .... that could provide additional information about why an observed phenomenon occurs” (Bennett, 2000).
- Child mental health “rapid progress ... depends on efforts to identify ... mediators of treatment outcome. We recommend randomized clinical trials routinely include and report such analyses” (Kraemer et al., 2002).



# Historical Precedent

- Stimulus>Organism >Response (SOR) theory:
  - the effect of a Stimulus on a Response depends on mechanisms in the organism (Woodworth, 1928). These mediating mechanisms translate the Stimulus to the Response. SOR theory is ubiquitous in psychology.
- Stimulus: Multiply 24 and 16
- Organism: You
- Response: Your Answer
- Organism as a Black Box



# Two-, Three-, or Four-Variable Effects

- Two variables
  - $X \rightarrow Y, Y \rightarrow X, X \rightarrow Y$  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, moderator/interaction.
- Four variables
  - many possible relations among variables, e.g.,  $X \rightarrow Z \rightarrow M \rightarrow Y$ .



# Mediator versus Confounder

- Confounder is a variable related to two variables of interest that falsely obscures or accentuates the relation between them (Meinert & Tonascia, 1986).
- The definition below is also true of a confounder because a confounder also accounts for the relation but it is not intermediate in a causal sequence.
- In general, a mediator is a variable that accounts for all or part of the relation between a predictor and an outcome (Baron & Kenny, 1986, p.1176).



# Mediator versus Moderator

- Moderator is a variable that affects the strength of the relation between two variables. The variable is not intermediate in the causal sequence so it is not a mediator.
- Moderator is usually an interaction; the relation between X and Y depends on a third variable. There are other more detailed definitions of a moderator.



# Reasons for Mediation Analysis in Intervention Research

- Mediation is important for health research. Practical implications include reduced cost and more effective interventions.
- Mediation analysis is based on theory for the processes underlying interventions.
  - Action theory corresponds to how the intervention will affect mediators. Treatment Theory (Piper et al., 2007)
  - Conceptual Theory focuses on how the mediators are related to the outcome variables (Chen, 1990, Lipsey, 1993). Relapse Theory (Piper et al., 2007)



# Questions about Mediators Selected for Intervention Research

- Are these the right mediators? Are they causally related to the outcome? Is self-esteem causally related to externalizing behavior? → Conceptual or Treatment Theory
- Can these mediators be changed? Can personality be changed? → Action Theory
- Will the change in these mediators that we can muster with our intervention be sufficient to lead to abstinence? Do we have the resources to change personality in four sessions? → Both Action and Conceptual Theory



# Audience Interaction

- “Now, could you come up with or describe mediators in your research?”

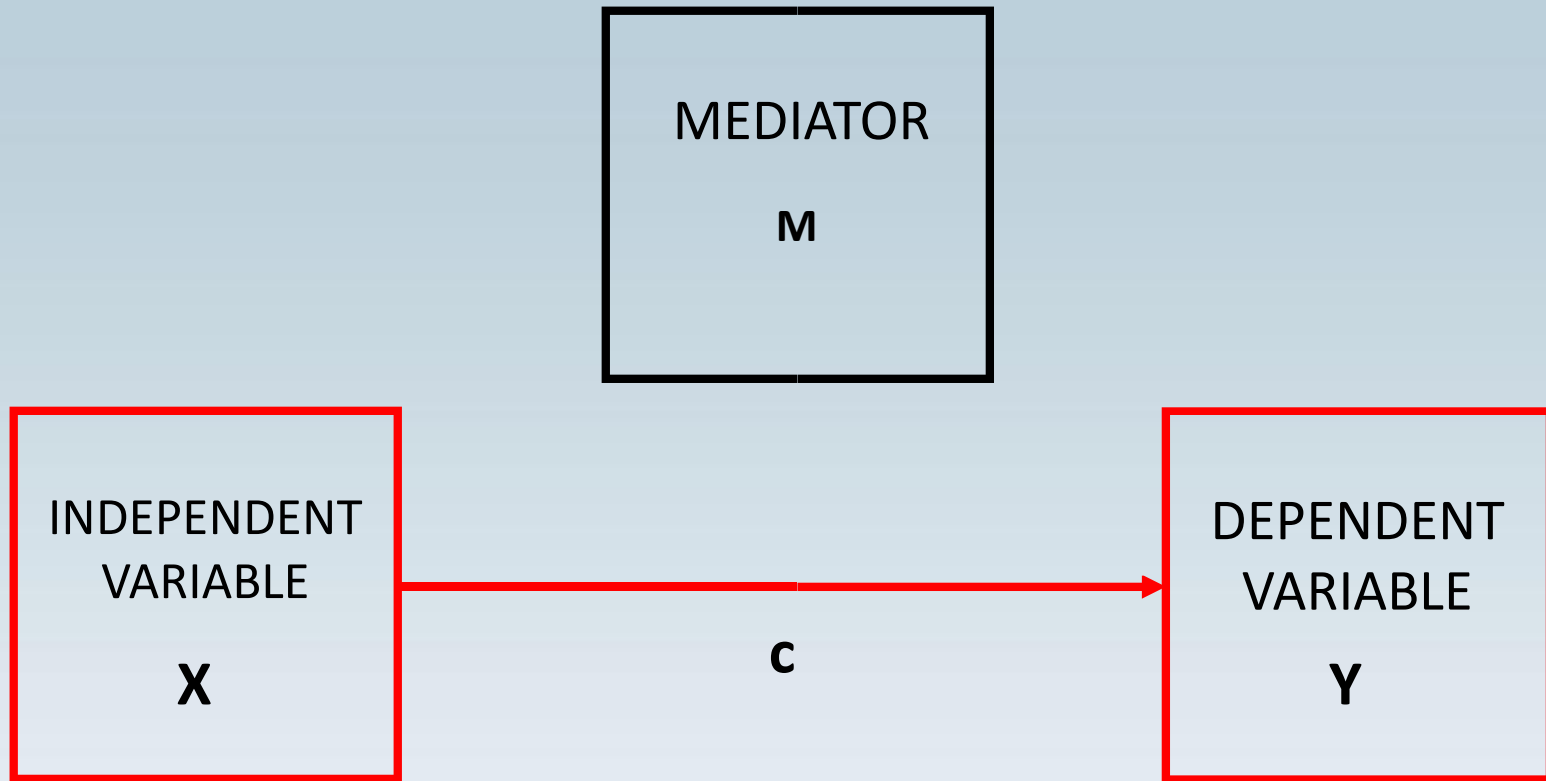


# Causal Steps Test of Mediation

- Series of steps described in Judd & Kenny (1981) and Baron & Kenny (1986)
- One of the most widely used methods to assess mediation in psychology
- Consists of a series of tests required for mediation as shown in the next slides



# Regression Equation 1

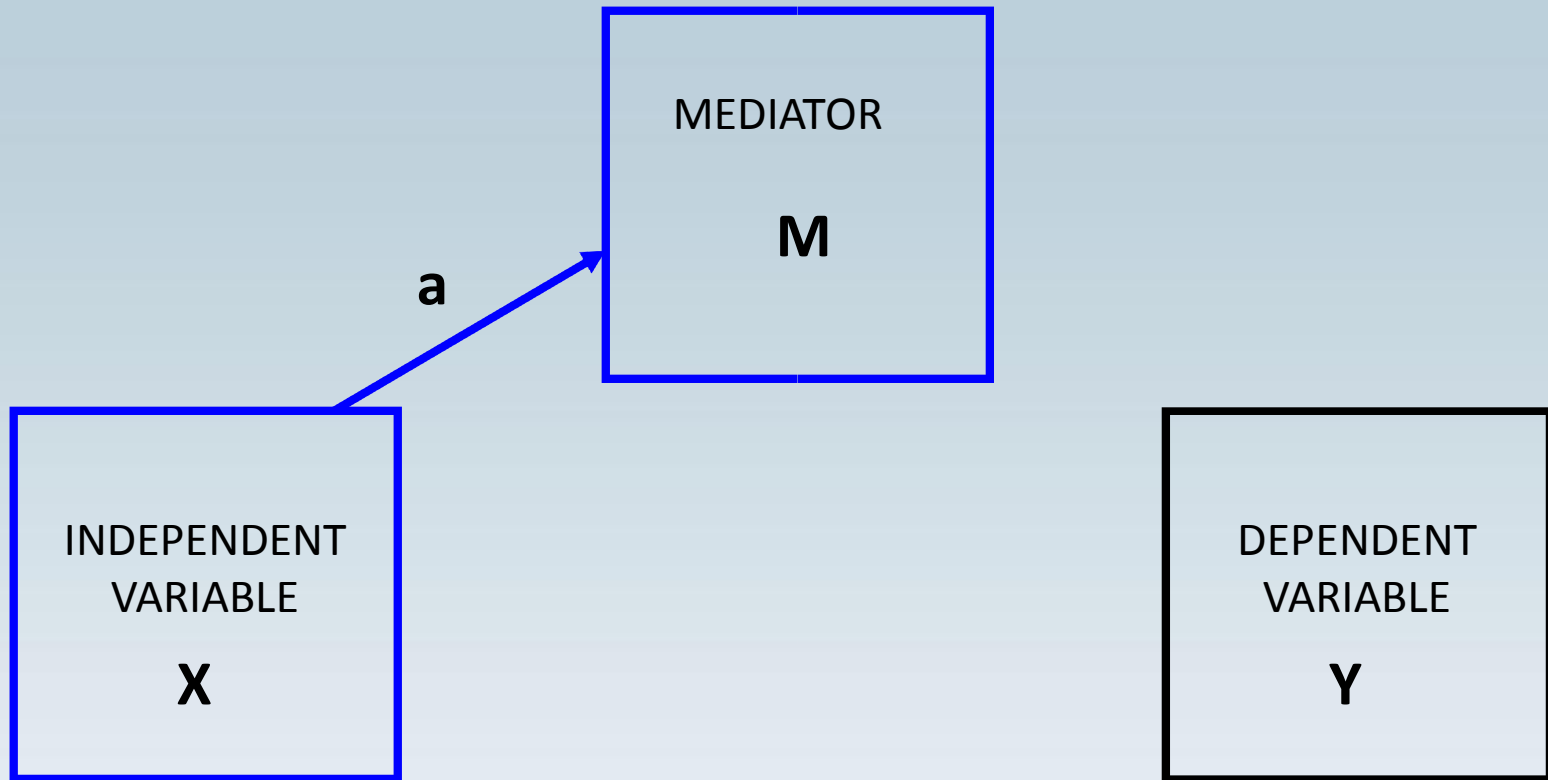


The independent variable is related to the dependent variable:

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



# Regression Equation 2

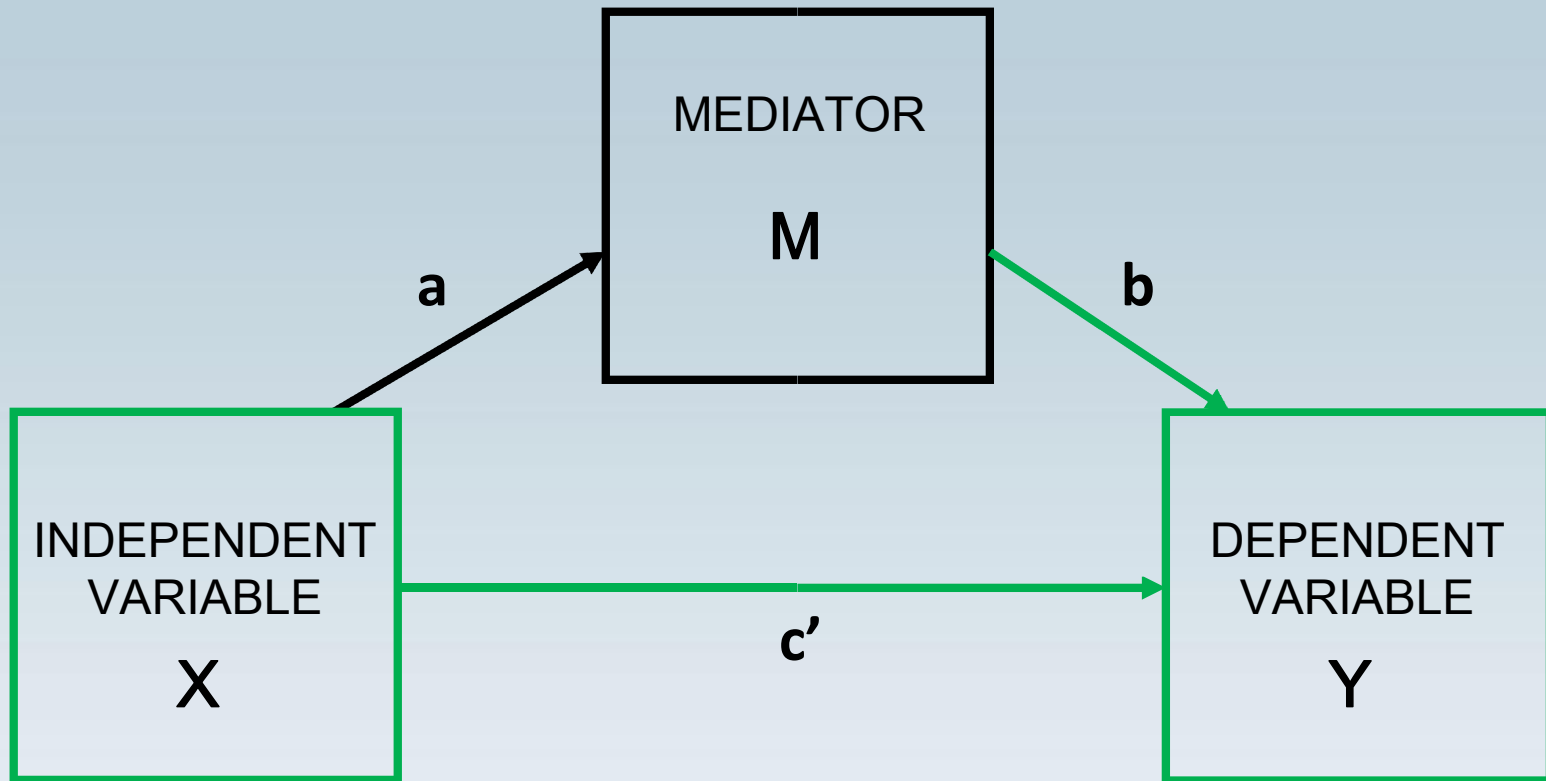


The independent variable is related to the potential mediator:

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



# Regression Equation 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= $c'$  and Total effect= $ab+c'=c$



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

- Mediated effect =  $\hat{a}\hat{b}$  Standard error =  $\sqrt{\hat{a}^2 s_b^2 + \hat{b}^2 s_a^2}$
- Multivariate delta method standard error (Sobel 1982; Folmer 1981)
- Test for significant mediation:
- $z' = \frac{\hat{a}\hat{b}}{\sqrt{\hat{a}^2 s_b^2 + \hat{b}^2 s_a^2}}$  Compare to empirical distribution of the mediated effect



# Assumptions

- For each method of estimating the mediated effect based on Equations 1 and 3 (c-c') or Equations 2 and 3 (ab):
  - Predictor variables are uncorrelated with the error in each equation.
  - Errors are uncorrelated across equations (ab).
  - Predictor variables in one equation are uncorrelated with the error in other equation.
- Reliable and valid measures
- No omitted influences
- Normally distributed variables



# Assumptions, cont.

- Data are a random sample from the population.
- Coefficients  $a$ ,  $b$ ,  $c'$  reflect true causal relations and the correct functional form.
- Mediation chain is correct: Temporal ordering is correct,  $X$  before  $M$  before  $Y$ . Any mediation model is part of a longer mediation chain. The researcher decides what part of the micromediation chain to examine.
- Homogeneous effects across subgroups: It is assumed that the relation from  $X$  to  $M$  and from  $M$  to  $Y$  are homogeneous across subgroups or other characteristics of participants in the study. This means there are not moderator effects.



# Reasons for Confidence Limits

- Gives a range of values based on a sample estimate.
- Helps avoid binary, significant or not, approach to research.
- Incorporates variability in the point estimate as well as the point estimate.



# Confidence Limits for $\hat{a}\hat{b}$

- Confidence Limits  $\hat{a}\hat{b} \pm z_{crit} s_{\hat{a}\hat{b}}$
- UCL =  $\hat{a}\hat{b} + z_{crit} s_{\hat{a}\hat{b}}$
- LCL =  $\hat{a}\hat{b} - z_{crit} s_{\hat{a}\hat{b}}$
- Where  $z_{crit}$  is the z critical value because the standard error is asymptotic. Valid to use t instead of z.
- 95% Confidence Limits
- UCL =  $\hat{a}\hat{b} + 1.96 s_{\hat{a}\hat{b}}$
- LCL =  $\hat{a}\hat{b} - 1.96 s_{\hat{a}\hat{b}}$
- With normal distribution upper and lower critical values have the same value but opposite sign, e.g., 1.96 and -1.96.



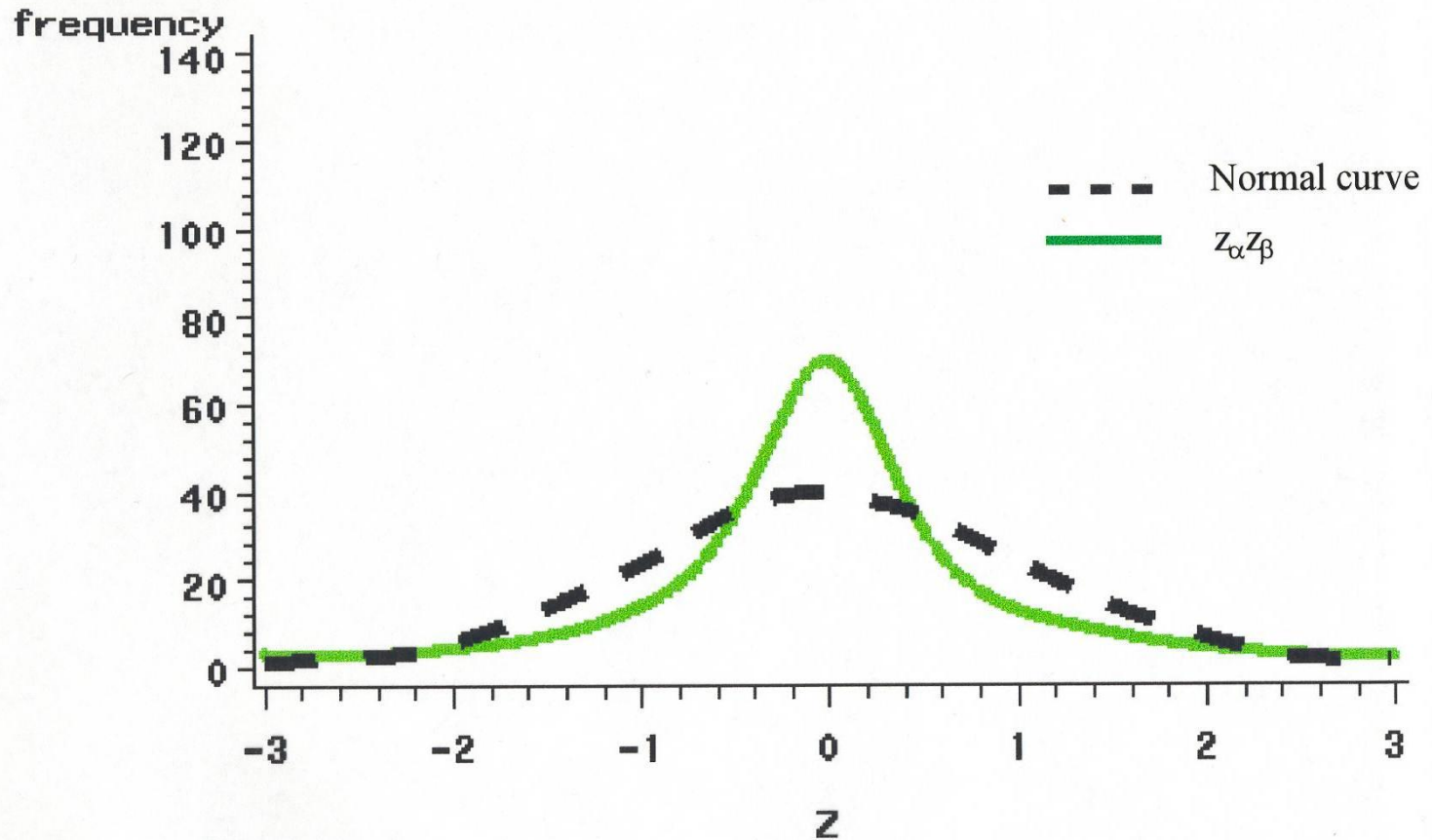
# Distribution of the Product

- The mediated effect is the product of two coefficients  $a$  and  $b$ . The distribution of the product has a normal distribution only in special cases.
- At low values of  $a$  and  $b$ , the distribution has excess kurtosis and skewness, e.g., when  $a$  and  $b$  are both zero, kurtosis is 6. It is not surprising that the confidence limits are inaccurate if the distribution is assumed to be normal.
- One solution is to use the distribution of the product in statistical tests and confidence limits.



# Distribution of $z_\alpha z_\beta$ vs. the normal curve

$\alpha\beta = 0, n=1000$



# PRODCLIN (PRODuct Confidence Limits for the INdirect effect)

- MacKinnon, Fritz, Williams, & Lockwood (2007) describes program to compute critical values for the distribution of the product.
- Web location includes programs in SAS, SPSS, and R that access a FORTRAN program.
  - <http://www.public.asu.edu/~davidpm/ripl/Prodclin/>
- Input estimates  $\hat{a}$ ,  $se_{\hat{a}}$ ,  $\hat{b}$ ,  $se_{\hat{b}}$ , correlation between  $\hat{a}$  and  $\hat{b}$ , and Type I error rate. Output includes the input values and normal and distribution of the product confidence limits.



# Three Major Types of Single Sample Tests for the Mediation Effect

1. Causal Steps: Series of tests described in Baron and Kenny (1986), for example
2. Difference in Coefficients:  $c - c'$ , e.g., from Clogg et al. (1992)
3. Product of Coefficients:  $ab$ , Sobel (1982) and distribution of the product

See MacKinnon et al., *Psychological Methods* (2002) for a review and comparison of single sample tests



# Empirical Sample Size Estimates for .8 Power to Detect the Mediated Effect

<u>Test</u>	<u>S-S</u>	<u>S-M</u>	<u>S-L</u>	<u>M-S</u>	<u>M-M</u>	<u>M-L</u>	<u>L-S</u>	<u>L-M</u>	<u>L-L</u>
<u>Baron/Kenny</u> (causal steps)	<b>20886</b>	3039	1561	2682	397	204	1184	175	92
<u>a &amp; b Joint</u> (causal steps)	530	403	403	405	74	58	405	59	36
<u>Sobel</u> (product test)	667	422	412	421	90	66	410	67	42
<u>PRODCLIN</u> (product test)	539	401	402	404	74	57	404	58	35

**Note:**  $c' = 0$  for the models,  
i.e., there was no direct effect. Two-tailed tests

(Fritz & MacKinnon, *Psychological Science*, 2007)



# Best Statistical Test for Mediation

- Investigate mediation whether or not the total effect is statistically significant.
  - Test of mediation can have more power and investigation of action and conceptual theory is important regardless of significance of the overall effect.
- Use a product of coefficients,  $ab$ , test such as the distribution of the product, bootstrap, or joint significance.
- Compute confidence intervals and effect size for the mediated effect.

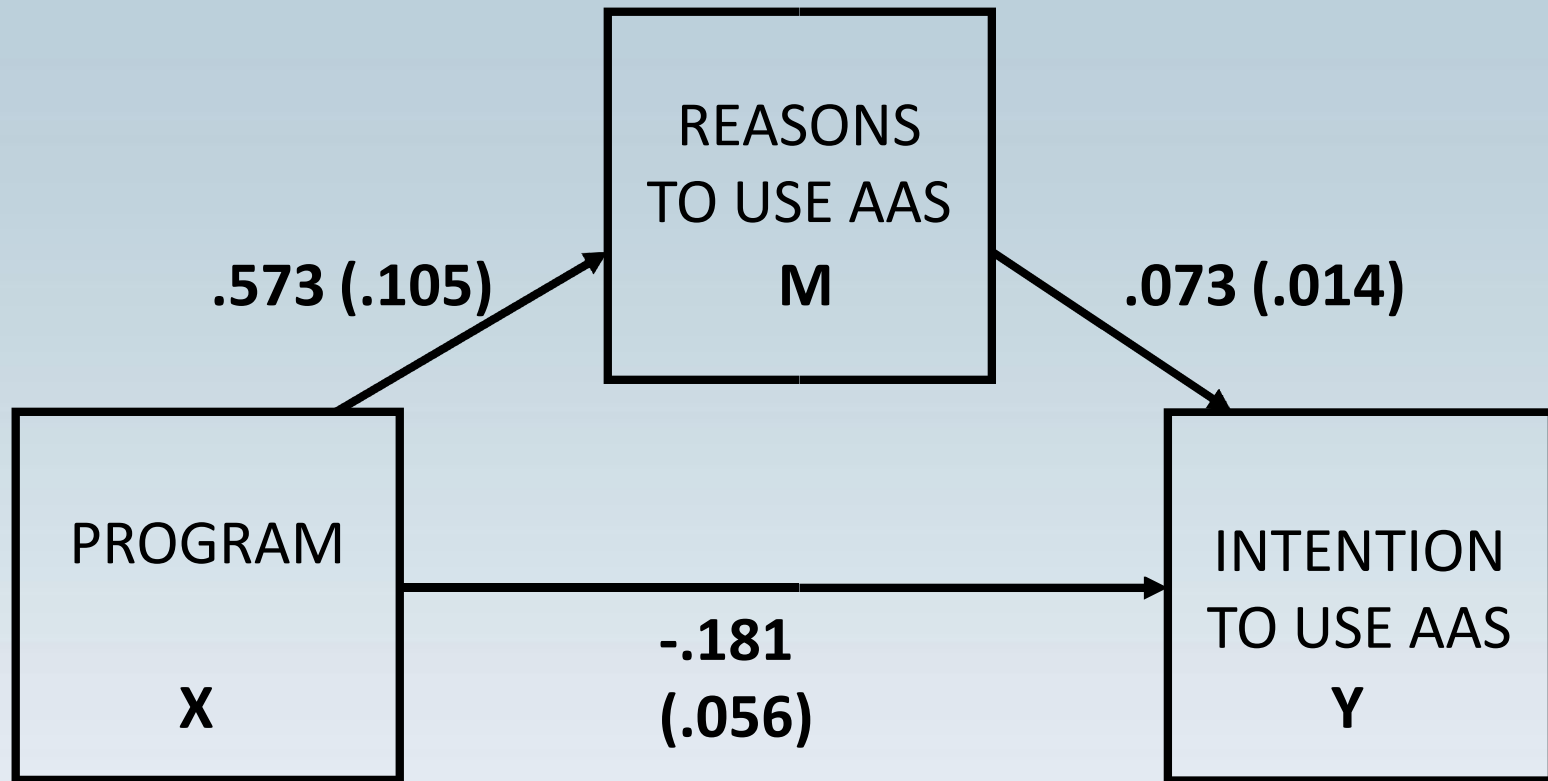


# Special Topic: Inconsistent Mediation Models

- Inconsistent mediation models: at least one of the mediated effects and direct effects have different signs (see MacKinnon, Krull, & Lockwood 2000).
- If the overall effect of X on Y is zero but there is a significant mediated effect, then it is an inconsistent mediation model. These effects are sometimes called suppressor effects. In these models the effect of X on Y actually increases when the mediator is included in the model.
- “one may be equally misled in assuming that an absence of relation between two variables is real, whereas it may be due...to the intrusion of a third variable” (Rosenberg, 1968, p. 84).



# Inconsistent Mediation in ATLAS Data

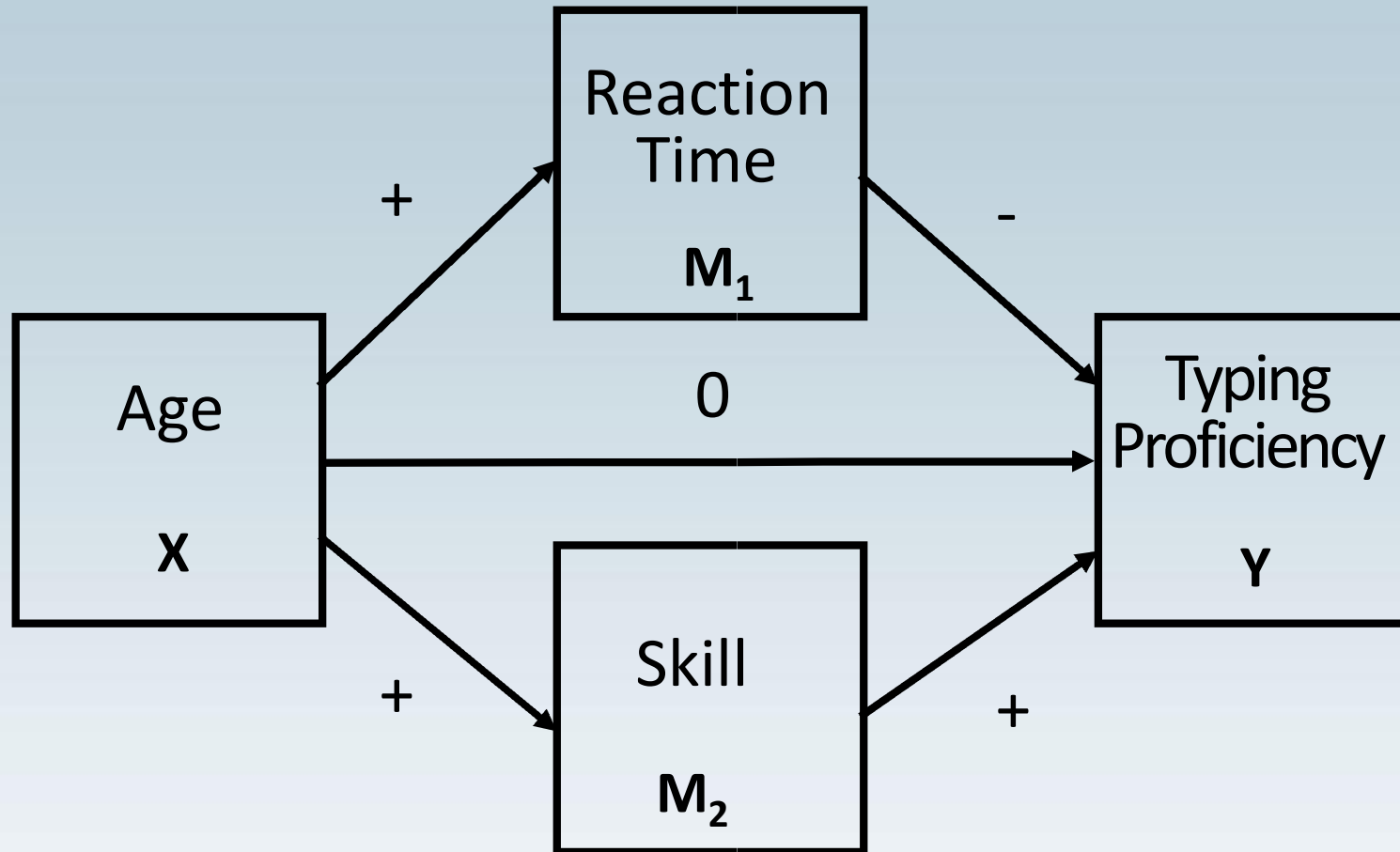


Mediated effect = .042

Standard error = .011



# Mediators of the Null Effect of Age on Typing (Salthouse, 1984)



# What if repeated measures of X, M, and Y are available?

- Measures of X, M, and Y at two time points allow for several options: difference score, ANCOVA, residualized change score.
- Measures of X, M, and Y at three or more time points allow for many alternative longitudinal models: Autoregressive, Latent Growth, Latent Difference Score Models, Survival Models.
- For treatment research, X is usually measured once and represents random assignment of participants to one of two groups.

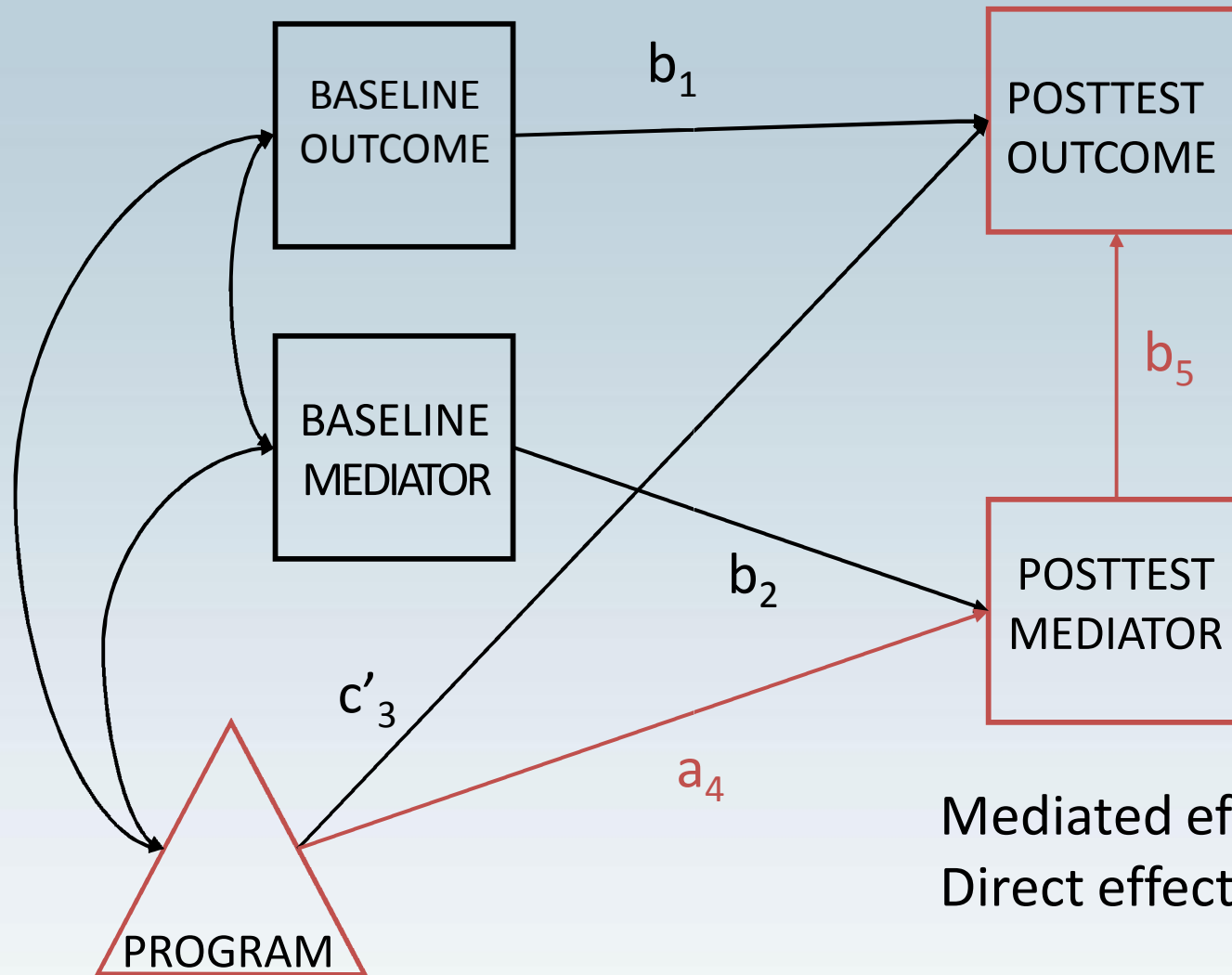


# Longitudinal Mediation Analysis

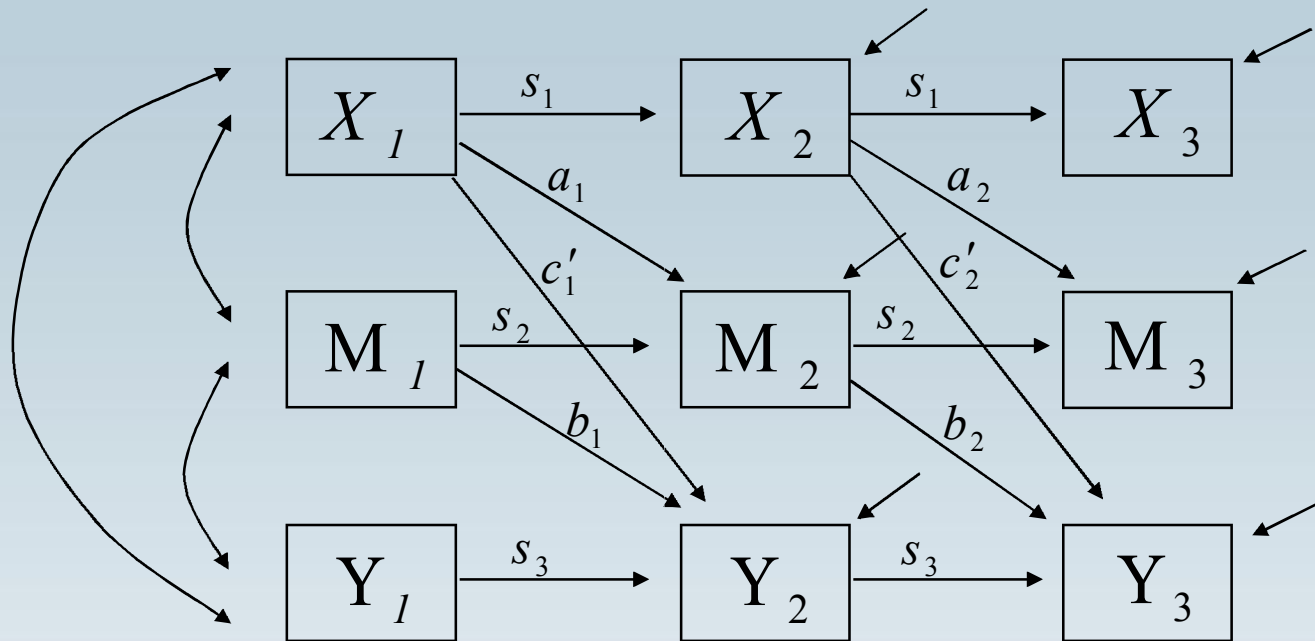
- Assume correct temporal ordering: X before M before Y.
- Relations among X, M, and Y are at some equilibrium so the observed relations are not solely due to when they are measured.
- Correct timing and spacing of measures to detect effects.
- Consider when X affects M and when M affects Y.
- Triggering, cascading, and other timing processes may be at work (Tang & DeRubeis, 1999; Howe et al., 2003).
- Timing is crucial for deciding when to collect longitudinal measures (Collins & Graham, 2003).



# Two-Wave Longitudinal Model



# Autoregressive Model with Time-Ordered Mediation



*Note: All residuals are correlated*

Cole & Maxwell (2003)



# Mediated Effects in the Autoregressive Model

## **Time-ordered** & **contemporaneous** mediation:

***a1b2***- path from X1 to M2 and path from M2 to Y3

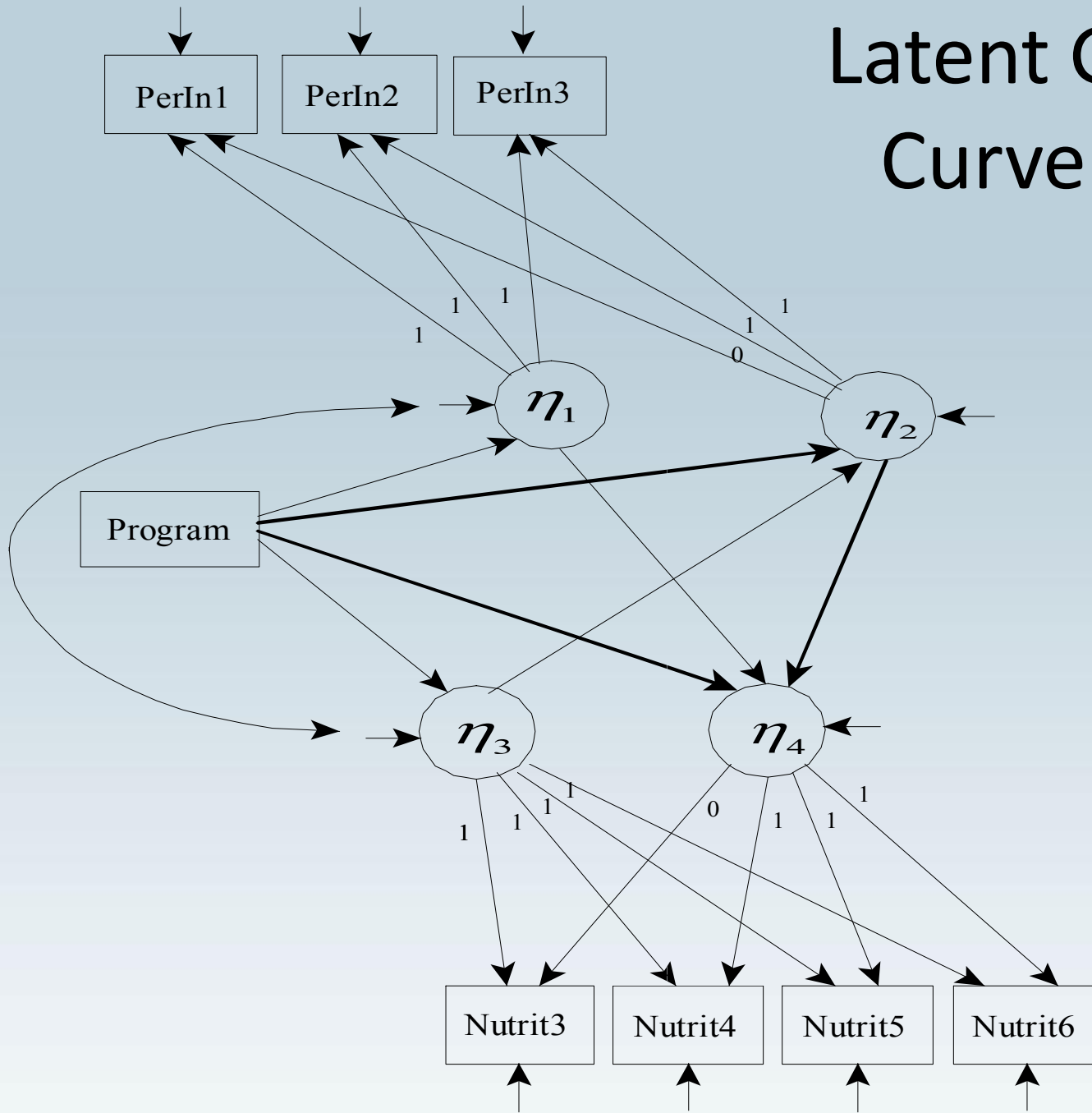
***a3b3***- path from X2 to M2 and path from M2 to Y2

***a4b4***- path from X3 to M3 and path from M3 to Y3

\*Note that model does not allow for random effects for individual change and does not include modeling of means. Change in growth of means is an important aspect of longitudinal data.



# Latent Growth Curve Model



# Survival Models

- Outcome is whether or not and when an event occurred.
- Discrete time survival models (Muthen & Masyn, 2004) are appropriate in many studies where time is measured discretely, e.g., alcohol status in each of the 3 weeks before and 5 weeks after a randomized treatment.
- Mplus computer program (Muthen & Muthen, 2004) can be used to estimate the parameters of this model.



# Causal Inference for Mediators

- Statistical approaches to improving causal inference from a mediation study:
  - I. Instrumental Variables, Holland, 1988; Sobel 2006.
  - II. Principal Stratification and latent classes; Frangakis & Rubin, 2002; Jo, 2008
  - III. G-estimation, Robins, 1999; Robins & Hernan, 2006, Greenland et al., 2008
- Exposes limitations of mediation analysis



# Design Approaches to Identifying Mediators

- Statistical mediation analysis asks: “How does a researcher use measures of the hypothetical intervening process to increase the amount of information from a research study?”
- Design approaches ask: “What is the best next study or studies to conduct after a statistical mediation analysis to further test the underlying mediation theory?”
- Five overlapping design approaches to indentifying mediators:
  - 1) Double Randomization
  - 2) Blockage
  - 3) Enhancement
  - 4) Purification
  - 5) Pattern matching for multiple variables, subgroups, settings, time, and alternative manipulations



# Blockage Designs

- The goal of blockage designs is to test a mediation relation with a manipulation that blocks the mediator from operating.
- Example: An exercise program appears to reduce depression perhaps by increasing social activity involved in exercising. A blockage manipulation would randomly assign participants to exercise alone so that persons would have reduced social activity. If social activity is the mediator of exercise effects on depression, then reduced effects would be observed among those exercising alone.



# Enhancement Designs

- The goal of enhancement designs is to deliver interventions that enhance the effects of a hypothesized mediator.
- Example: An addiction treatment program reduces remission by improving social support. An enhancement design would increase social support even more to demonstrate a larger effect on remission. Social support may be increased by more exposure to a therapist, additional contact with friends and family, etc.

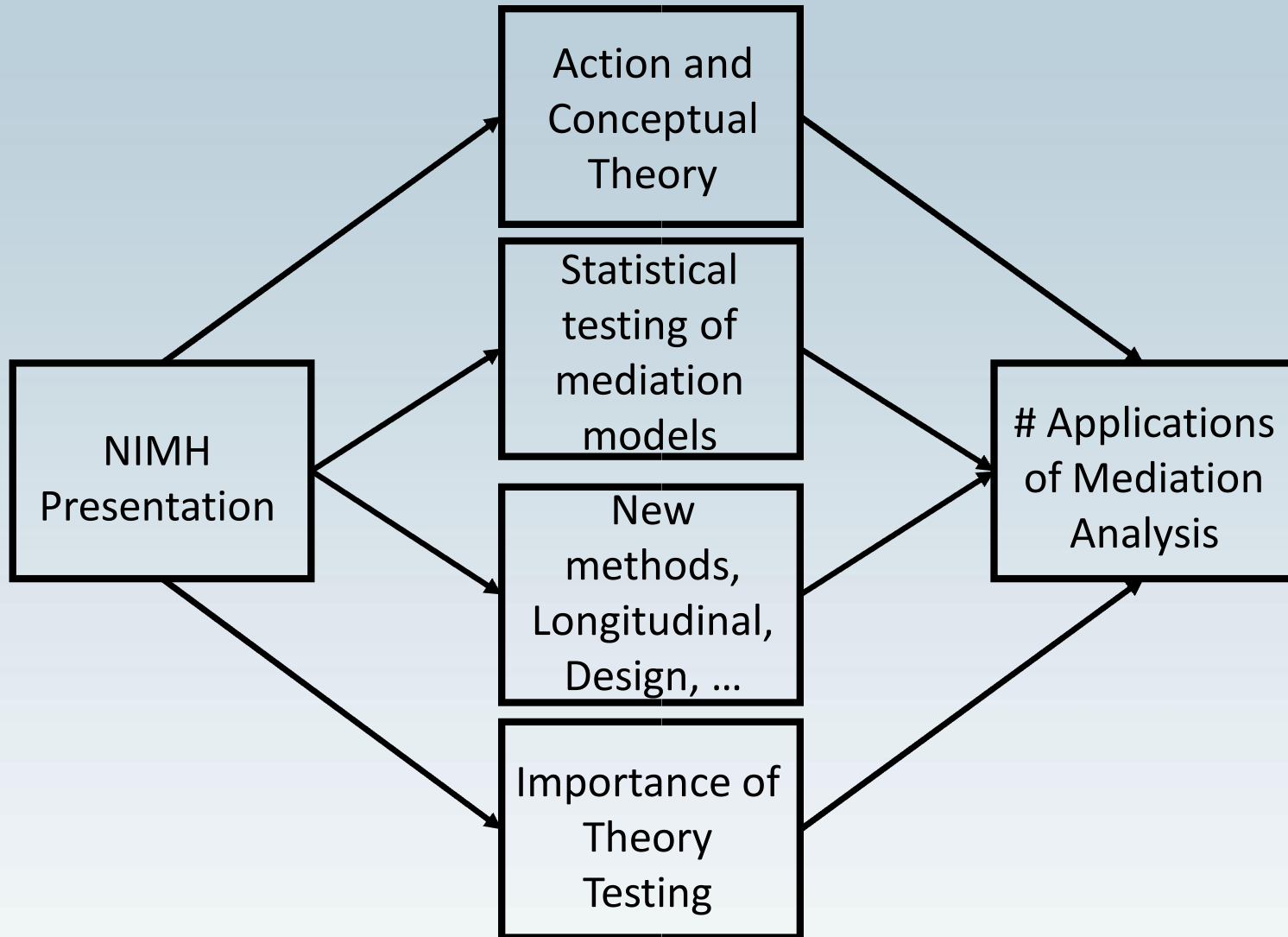


# Summary

- Mediation analysis is important because it provides information on how two variables are related or how an intervention achieved its effects.
- Tests of mediation based on the product of coefficients are the best method.
- Longitudinal data analyses provide an ideal way to test for mediation.
- Experimental designs targeting critical mediators help further clarify the mediating mechanisms surrounding the most effective treatments.



# Hypothesized Effects of NIMH Akron Mediation Analysis Presentation



Dissemination and networking  
services brought to you by



**interlink**<sup>TM</sup>  
training dissemination services

**interlink***your* **training**.com