ADVANCES IN WITHIN-SUBJECT MEDIATION ANALYSIS

Andrew F. Hayes and Kristopher J. Preacher, co-chairs

Authors and Presenters

Andrew F. Hayes*, Ph.D., The Ohio State University

Amanda K. Montoya*, The Ohio State University

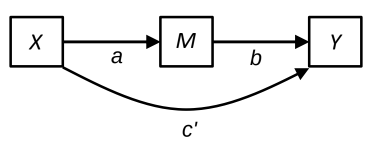
Kristopher J. Preacher*, Ph.D, Vanderbilt University

James P. Selig, Ph.D., University of Arkansas

Elizabeth Page-Gould, Ph.D., University of Toronto

Amanda Sharples*, University of Toronto

* presenter



- In this symposium, we each address statistical approaches to mediation analysis in studies that involve repeated measurement of *X*, *M*, or *Y* rather than merely observed or manipulated cross-sectionally and measured only once.

Specifically, we address…

- an empirical approach to examining the problem of determining the effect of measurement lag on indirect effects (Preacher and Selig).
- a path-analytic approach to quantifying and testing indirect effects in the two-condition experiment where *M* and *Y* are repeatedly measured in people assigned to **both** conditions of *X* (Hayes and Montoya).
- multilevel random-effects mediation models when *X*, *M*, and *Y* are repeatedly measured on the same person using a variety of stimuli or in a variety of situations (Page-Gould and Sharples).

In the interest of time, please save your questions for the Q&A period.

2016 SPSP ANNUAL CONVENTION #SPSP2016

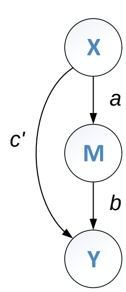
STRATEGIES FOR INCORPORATING LAG AS MODERATOR IN MEDIATION MODELS

Kristopher J. Preacher
Vanderbilt University

James P. Selig
University of Arkansas for Medical Sciences

The common "simple mediation" model

Consider the simple mediation model commonly used in social and personality psychology studies:



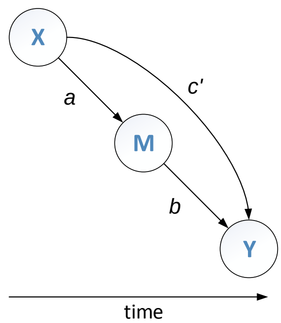
This model has great heuristic value. Yet methodologists have had much to say about the inadequacy of this model for drawing causal conclusions.

Perhaps the biggest limitation is that many designs assess *X*, *M*, and *Y* simultaneously, or nearly so.

Yet, causes need time to exert their effects (Hume, 1738).

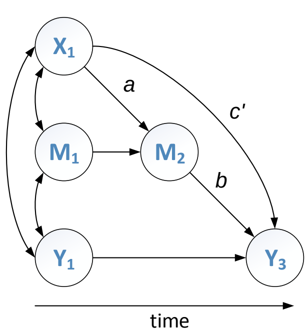
2016 SPSP ANNUAL CONVENTION #SPSP2016 SP SP

With measurement staggered in time



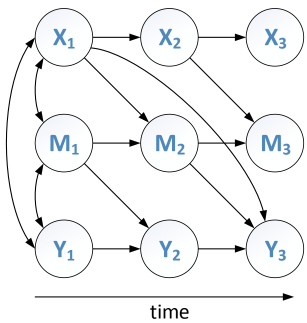
We could stagger the assessments of *X*, *M*, and *Y* in time, allowing time for the effect and to be more confident when saying things like "*X* causes *Y* indirectly through *M*" or "*M* mediates the effect of *X* on *Y*."

Adding covariance adjustment for prior measurements



Controlling for prior measurements of *M* and *Y* is also recommended. This helps separate out the stable variance in *M* and *Y*, which cannot be explained by other predictors.

The cross-lagged panel mediation model



Combining these recommendations leads to the popular cross-lagged panel model (CLPM) approach to assessing mediation (Cole & Maxwell, 2003).

Effects are not invariant to choice of time lag

The CLPM is a more defensible method for assessing mediation. However, it still suffers from a major problem—the effects in such models *depend on the chosen lag*, or how much time elapses between the assessments of *X*, *M*, and *Y*.

From Voelkle et al. (2012):

Two researchers studying the same variables use two different lags (1 vs. 2 mos.) and draw different conclusions about the strength of the *X*→*Y* and *Y*→*X* effects.

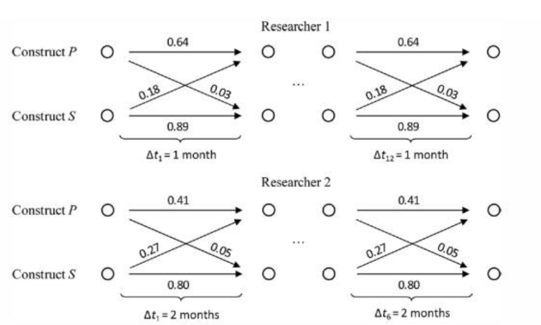
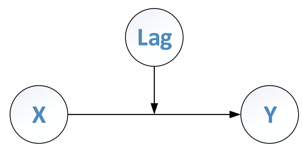


Figure 1. Autoregressive and cross-lagged parameter estimates of two studies on the relationship between two constructs across $T = 12$ versus $T = 6$ intervals. All parameter estimates were constrained to equality over time, and time intervals are assumed to be of equal length within each study ($\Delta t_1, \ldots, \Delta t_{12} = 1$ month in Study 1, in the upper half of the figure, and $\Delta t_1, \ldots, \Delta t_6 = 2$ months in Study 2, in the lower half).

Lag as moderator (LAM) analysis

In the context of regression (*X*→*Y*), Selig, Preacher, & Little (2012) proposed using a *variable-lag design*, such that the assessment of either *X* or *Y* (or both) are deliberately staggered over time allowing lags to vary across persons.

The result is a *lag as moderator* (LAM) analysis, in which we explicitly model how the *X*→*Y* effect changes as a function of lag. Lag itself is treated as a moderator.



This can yield greater insight into the causal process, and can explain why different researchers arrive at different conclusions as a function of the arbitrary amount of time that elapses between the assessment of different variables.

Lag as moderator (LAM) analysis

Here is how it works. Rather than using: $\hat{Y}_i = b_0 + b_1 X_i$

We proposed instead using:

$$\begin{aligned}\hat{Y}_i &= b_0 + b_1 X_i + b_2 Lag_i + b_3 X_i Lag_i \\ &= \left(b_0 + b_2 Lag_i\right) + \left(b_1 + b_3 Lag_i\right) X_i\end{aligned}$$

...an example of a standard interaction model, where $\left(b_1 + b_3 Lag_i\right)$ is the *simple slope* relating *X* to *Y* at a given lag.

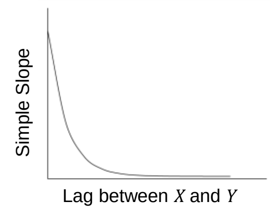
This model requires individual differences in lag, which can be either observational or experimentally manipulated.

Allowing for nonlinearity in the effect of lag

The previous model assumes that the effect of *X* on *Y* varies as a *linear* function of lag, which may be approximately true in many cases.

But Selig et al. discuss nonlinear alternatives that may be more realistic in a given setting. For example, the effect of *X* on *Y* may follow a *negative exponential* function of lag:

$$\hat{Y}_i = b_0 + b_1 e^{b_2 Lag_i} X_i$$



Estimation requires nonlinear regression, but it can be done with programs like SAS, SPSS, and R.

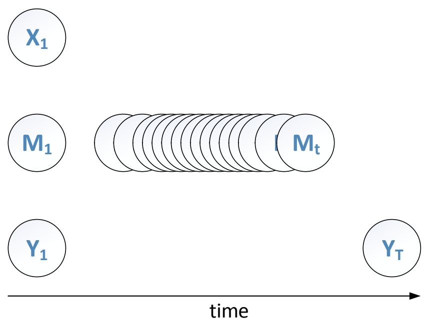
EMERALD

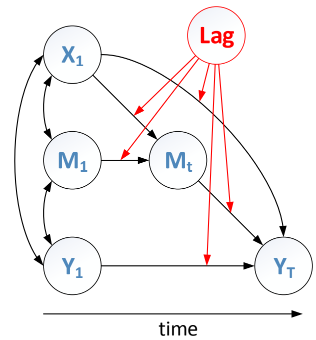
We propose extending the LAM approach to mediation analysis, an approach we term *Examining Mediation Effects using a Randomly Assigned Lags Design* (EMERALD).

Using the EMERALD, researchers would *deliberately vary the lags* separating assessments, then incorporate lag into the model as a moderator of the mediation paths *a* and/or *b*.

EMERALD

For example, one could hold *X* and *Y* fixed in time, deliberately stagger the assessment of *M*, and estimate the *a* and *b* slopes of a mediation model conditional on lag.





Or, one could hold *M* fixed in time and stagger the assessment of *X* and *Y*, etc.

EMERALD

It is straightforward to use ordinary SEM for LAM with linear moderation by lag.

If the moderation by lag is nonlinear (e.g., exponential), could use "constraint variables" in Mplus or "definition variables" in Mx—still in the SEM framework.

An Illustration: Measurement

We used data from a large longitudinal prevention study (Goldberg et al., 1996) to illustrate a simple application of the EMERALD with *X* and *Y* fixed in time, and the timing of *M* allowed to vary.

X: Intervention Status (Program = 0; Control = 1) was randomly assigned at the beginning of the study.

M: Beliefs about the Severity of Steroid Use was assessed on one of three occasions: approximately 0, 2, and 12 months after the beginning of the study.

Y: Intention to Use Steroids was assessed once approximately 14 months after the beginning of the study.

Goldberg, L., Elliot, D., Clarke, G.N., MacKinnon, D.P., Moe, E., Zoref, L., Green, C., Wolf, S.L., Greffrath, E., Miller, D.J. & Lapin, A., 1996. Effects of a multidimensional anabolic steroid prevention intervention: The Adolescents Training and Learning to Avoid Steroids (ATLAS) Program. *JAMA*, 276(19), pp.1555-1562.

An Illustration: Data extraction for EMERALD

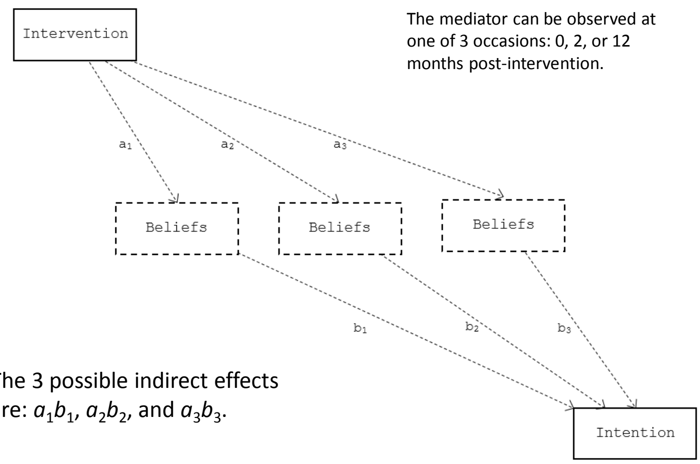
Using the fully longitudinal data, we created an EMERALD study with each participant having one observed value for *X*, *M*, and *Y*.

Times of measurement for *X* and *Y* were the same for all participants.

The value for the mediator was randomly selected from values at three different occasions (0, 2, or 12 months after the study began).

An Illustration: Indirect effects as a function of lag

EMERALD Design



The mediator can be observed at one of 3 occasions: 0, 2, or 12 months post-intervention.

The 3 possible indirect effects are: a_1b_1, a_2b_2, and a_3b_3.

Estimation and inference

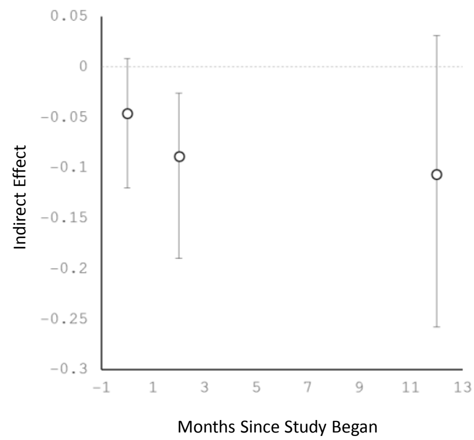
With only three discrete lag values, we chose a multi-group regression analysis to separately estimate the indirect effect at the three different values of lag.

We computed 95% confidence intervals for the indirect effect using a Monte Carlo strategy (Preacher & Selig, 2012).

2016 SPSP ANNUAL CONVENTION #SPSP2016

Results

Indirect Effects and 95% Confidence Intervals across Three Lags

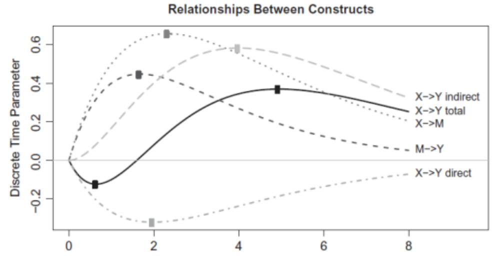


In continuous time

CLPM: Longitudinal, and easy to apply, but reflects a “snapshot” of mediation at only a single, arbitrary lag.

EMERALD: Yields indirect effects as a function of lag, but requires collecting data such that lag varies across persons. Can be fit in any SEM program.

Deboeck & Preacher (2016) describe *continuous time mediation models*. These models require data collected at only one choice of lag, but yield indirect effects at any chosen lag. Requires differential equations and specialized software.



Take home points

- Many effects will vary with lag, yet lags are often chosen arbitrarily.
- Failures to replicate results may be due to varying lags between studies.
- Everyone should record variability in lag, whether observed or manipulated.
- It is feasible to study lag-dependent effects. We have options now!

ESTIMATION AND INFERENCE ABOUT INDIRECT EFFECTS IN WITHIN-SUBJECTS MEDIATION ANALYSIS: A PATH ANALYTIC PERSPECTIVE

Andrew F. Hayes, Ph.D.
The Ohio State University

ESTIMATING AND COMPARING INDIRECT EFFECTS IN TWO-CONDITION WITHIN-SUBJECT MULTIPLE MEDIATOR MODELS

Amanda K. Montoya
The Ohio State University

Based on Montoya, A. K., & Hayes, A. F. (2015). Two-condition within-participant statistical mediation analysis: A path-analytic framework. In review (first R&R) at *Psychological Methods*

The many details skipped due to time constraints are available in the paper, downloadable from the Mechanisms and Contingencies Lab web page at www.afhayes.com

An exemplar of the common two-condition within-subject experimental design

Data are from Dohle, S., & Siegrist, M. (2014). Fluency of pharmaceutical drug names predicts perceived hazardousness, assumed side effects, and willingness to buy. *Journal of Health Psychology*, 19, 1241-1249.

22 participants presented with the names of 10 drugs, 5 with simple names (e.g., Fastinorbine), and 5 with complex names (e.g,. Cyrigmcmium).

M = Perceived hazardousness (1 to 7, higher = more)
Y = Willingness to purchase (1 to 7, higher = more)
Measurement 1 = Average judgment about drugs with simple names
Measurement 2 = Average judgment about drugs with complex name

	Simple		Complex	
ID	M_1	Y_1	M_2	Y_2
1	3.8	4.4	4.4	3.6
2	4.2	4.2	5.2	2.0
3	4.0	4.0	4.0	4.0
4	4.4	3.0	3.0	5.2
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
22	3.2	4.2	5.8	2.8
Mean	3.9	3.9	4.7	3.3

Analytical goal: Determine if perceived hazardousness of the drug is a *mediator* of the effect of the drug name complexity on willingness to purchase.

Judd, Kenny, and McClelland (2001)

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychological Methods*, *6*, 115-134.


Estimating and Testing Mediation and Moderation in Within-Subject Designs

Charles M. Judd
University of Colorado at Boulder

David A. Kenny
University of Connecticut

Gary H. McClelland
University of Colorado at Boulder

One of the few treatments of mediation analysis in this common research design.

A “causal steps”, Baron and Kenny type logic to determining whether M is functioning as a mediator of X’s effect on Y when both M and Y are measured twice in difference circumstances but on the same people.

Judd et al.'s criteria to establish mediation

Analytical goal: Determine if perceived hazardousness of the drug is a *mediator* of the effect of the drug name complexity on willingness to purchase.

(1) Is there a difference between the two drugs types in participants' willingness to buy?

Yes, by a paired samples *t*-test.

(2) Is there a difference between the two drugs types in perceived hazardousness of the drug?

Yes, by a paired samples *t*-test.

(3) Does the difference in perceived hazardousness predict the difference in willingness to buy?

Yes, by a regression analysis.

(4) Does the difference in perceived hazardousness account for the difference in willingness to buy?

Yes, by a regression analysis. The difference in willingness to buy goes away when controlling for the difference in perceived hazardousness

	Simple		Complex	
ID	M_1	Y_1	M_2	Y_2
1	3.8	4.4	4.4	3.6
2	4.2	4.2	5.2	2.0
3	4.0	4.0	4.0	4.0
4	4.4	3.0	3.0	5.2
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
22	3.2	4.2	5.8	2.8
Mean	3.9	3.9	4.7	3.3

2016 SPSP ANNUAL CONVENTION #SPSP2016

Observations

(1) It seems very foreign relative to path-analytic approaches that now dominate mediation analysis in the between-subjects case. Where's the path analysis?

(2) This method is squarely rooted in the causal steps tradition to mediation analysis that has been much criticized. Compare it to the "Baron and Kenny" criteria:

- Is Y_2 statistically different than Y_1? This is like asking whether there is a total effect of *X* (drug came complexity) on *Y* (willingness to buy).
- Is M_2 statistically different than M_1? This is like asking whether *X* affects the mediator.
- Does difference in *M* significantly predict difference in *Y*? This is like asking whether the mediator affects the outcome.
- Is there still evidence of a difference in *Y* after accounting for the mediator? This is like asking whether the mediator completely or partially accounts for the effect of *X* on *Y*.

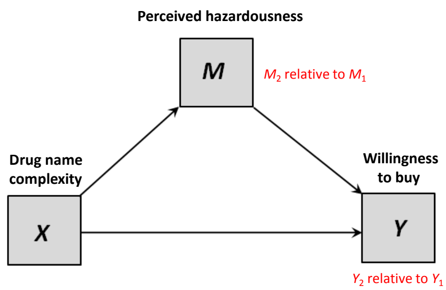
(3) There is no explicit quantification of the indirect effect, but it is the indirect effect that is the primary focus in 21st century mediation analysis.

All these things can be "fixed" by recasting JK&McC in a more familiar path-analytic form.

2016 SPSP ANNUAL CONVENTION #SPSP2016

In a path analytic mediation framework

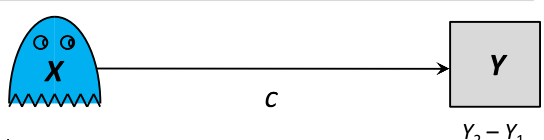
Goal: Model the effect of the drug name complexity on willingness to buy, directly as well as indirectly through the effect of the drug name complexity on perceived hazardousness.



	Simple		Complex	
ID	M_1	Y_1	M_2	Y_2
1	3.8	4.4	4.4	3.6
2	4.2	4.2	5.2	2.0
3	4.0	4.0	4.0	4.0
4	4.4	3.0	3.0	5.2
.	.	.	.	.
.	.	.	.	.
.				.
22	3.2	4.2	5.8	2.8
Mean	3.9	3.9	4.7	3.3

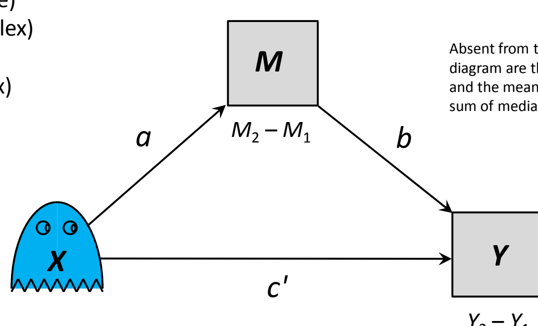
Where is *X* in the data?

In a path analytic mediation framework



Y_1 = willingness to buy (simple)
Y_2 = willingness to buy (complex)
M_1 = hazardousness (simple)
M_2 = hazardousness (complex)

Absent from this diagram are the errors and the mean centered sum of mediator values



Regression on just a constant.

$$Y_2 - Y_1 = c + e_1$$
$$M_2 - M_1 = a + e_2$$

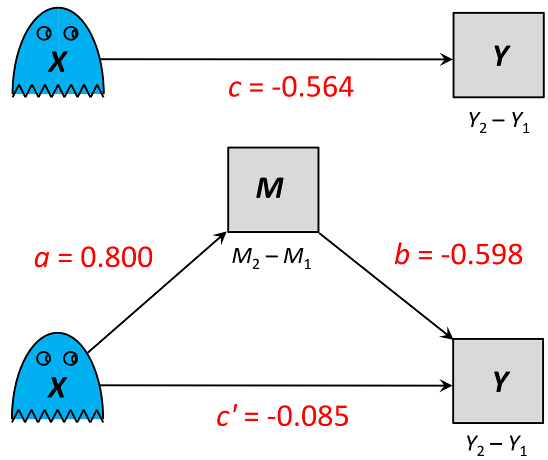
$$Y_2 - Y_1 = c' + b(M_2 - M_1) + b_2(M_2 + M_1)^* + e_3$$

* mean centered

$c = c' + ab$ $\quad$ $ab = c - c'$

In a path analytic mediation framework

In this form, it is clear that the effect of *X* partitions into two components direct and indirect in the usual way. We can conduct inferential tests on these estimates as in any mediation analysis.



$c = c' + ab$

$$c = \underbrace{-0.085}_{\text{Direct effect}} + \underbrace{(0.800)(-0.598)}_{\text{Indirect effect}} = \underbrace{-0.085}_{\text{Direct effect}} + \underbrace{-0.479}_{\text{Indirect effect}} = \underbrace{-0.564}_{\text{Total effect}}$$

Statistical inference for the indirect effect

What really matters in mediation analysis is the indirect effect *ab*. Some options include:

<u>“Sobel” test:</u>
Z = ab/se(ab), with p or confidence interval calculated assuming ab is normally distributed. This is not recommended because the sampling distribution of ab is not normal.

<u>Test of joint significance</u>
Are both a and b statistically significant? This is what Judd, Kenny, and McClelland use. We don’t recommend this as it requires two tests rather than one, and no interval estimate is provided.

<u>Monte Carlo confidence interval</u>
Assumes a normal sampling distribution of a and b individually, then simulates the sampling distribution of the product using Monte Carlo methods. This method is available in between-subjects mediation analysis and easy to do with the right software.

<u>Bootstrap confidence interval</u>
A natural choice as it assumes nothing about the sampling distribution of *ab*, and this is already common in between-subjects mediation analysis and easy to do with the right software.

Implementation: Mplus, PROCESS, and MEMORE

MPLUS

See handout or Montoya and Hayes (2015) for code and output.

PROCESS

PROCESS for SPSS and SAS (www.processmacro.org) can do this. How so is described in Montoya and Hayes (2015). See the discussion there.

MEMORE

MEMORE (**ME**diation and **MO**deration for **RE**peated measures; pronounced like "memory") is a bit easier to use than PROCESS for this kind of analysis but has PROCESS-like output. It is a new "macro" available for SPSS and SAS downloadable from *www.afhayes.com* and described for mediation problems in Montoya and Hayes (2015).

- Single and multiple mediator models.
- Various inferential methods for indirect effects
- Contrasts between indirect effects in multiple mediator models
- Moderated mediation analysis functions coming soon.

SPSS:
```
memore y=buy2 buy1/m=hazard2 hazard1/samples=10000.
```

SAS:
```
%memore (data=drugname,y=buy2 buy1,m=hazard2 hazard1,samples=10000);
```


MEMORE Output

```
************************ MEMORE Procedure for SPSS ***************************

                          Written by Amanda Montoya

                    Documentation available at afhayes.com

******************************************************************************

Variables:
Y = buy2     buy1
M = hazard2  hazard1

Computed Variables:
Ydiff =          buy2      -       buy1
Mdiff =          hazard2   -       hazard1
Mavg =   (       hazard2   +       hazard1  )       /2        Centered

Sample Size:
  22

******************************************************************************
Outcome: Ydiff =  buy2      -       buy1

Model
        Effect        SE         t        df         p      LLCI      ULCI
'X'     -.5636     .1932   -2.9168   21.0000     .0082    -.9655    -.1618

******************************************************************************
Outcome: Mdiff =  hazard2   -       hazard1

Model
        Effect        SE         t        df         p      LLCI      ULCI
'X'      .8000     .2579    3.1024   21.0000     .0054     .2637    1.3363

******************************************************************************
```

MEMORE constructs differences and averages for you.

c = -0.564

a = 0.800

MEMORE Output

```
*********************************************************************************
Outcome: Ydiff =  buy2      -       buy1

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .7721      .5961      .3667    14.0213     2.0000    19.0000      .0002

Model
              coeff         SE          t         df          p       LLCI       ULCI
'X'          -.0851      .1577     -.5399    19.0000      .5955     -.4152      .2449
Mdiff        -.5981      .1131    -5.2869    19.0000      .0000     -.8349     -.3613
Mavg         -.1818      .1683    -1.0803    19.0000      .2935     -.5341      .1705

********************** TOTAL, DIRECT, AND INDIRECT EFFECTS **********************

Total effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
     -.5636      .1932    -2.9168    21.0000      .0082     -.9655     -.1618

Direct effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
     -.0851      .1577     -.5399    19.0000      .5955     -.4152      .2449

Indirect Effect of X on Y through M
          Effect     BootSE   BootLLCI   BootULCI
Ind1      -.4785      .1363     -.7423     -.2063

Indirect Key
Ind1  X       ->      M1diff    ->       Ydiff

************************* ANALYSIS NOTES AND WARNINGS **************************

Bootstrap confidence interval method used: Percentile bootstrap.
Number of bootstrap samples for bootstrap confidence intervals:  10000
```

$c' = -0.085$

$b = -0.598$

$c = -0.564$

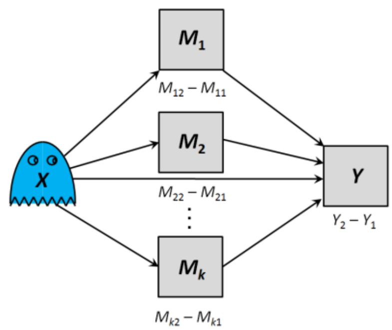
$c' = -0.085$

ab with 95% bootstrap confidence interval. This is consistent with a claim of mediation.

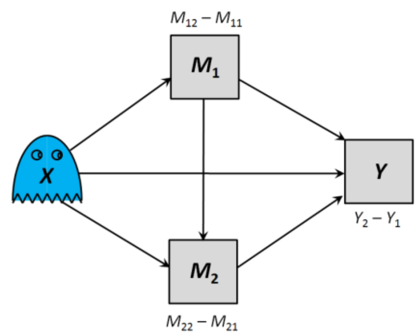
2016 SPSP ANNUAL CONVENTION #SPSP2016

Extension to multiple mediator models

A parallel multiple mediator model with k mediators



A serial multiple mediator model with 2 mediators



Why do this?

(1) More consistent with the complexity of real world-processes and theory.

(2) Allows for the testing of competing theories through different processes, as indirect effects can be formally compared.

An additional mediator measured in each condition

Data are still from Dohle, S., & Siegrist, M. (2014). Fluency of pharmaceutical drug names predicts perceived hazardousness, assumed side effects, and willingness to buy. *Journal of Health Psychology*, *19*, 1241-1249.

Participants also evaluated how effective they thought the drug would be.

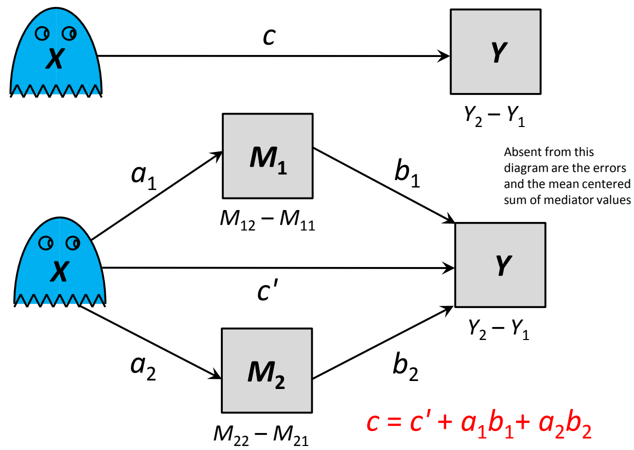
M_1 = Perceived hazardousness (1 to 7, higher = more)
M_2 = Perceived effectiveness (1 to 7, higher = more)
Y = Willingness to purchase (1 to 7, higher = more)
Measurement 1 = Average judgment about drugs with simple names
Measurement 2 = Average judgment about drugs with complex name

	Simple			Complex		
	M_{11}	M_{21}	Y_1	M_{21}	M_{22}	Y_2
	3.8	4.2	4.4	4.4	4.0	3.6
	4.2	4.4	4.2	5.2	3.6	2.0
	4.0	4.0	4.0	4.0	4.0	4.0
	4.4	4.2	3.0	3.0	4.8	5.2
	.	.	.	.	.	
	.	.	.	.	.	
	.	.	.	.	.	
	3.2	4.6	4.2	5.8	5.6	2.8
Mean	3.9	4.4	3.9	4.7	4.1	3.3

Analytical goal: Is the effect of drug name complexity on willingness to purchase mediated by hazardousness? effectiveness? Both? Are the indirect effects the same or different?

A parallel mediation model in path analytic form

Y_1 = willingness to buy (simple)
Y_2 = willingness to buy (complex)
M_{11} = hazardousness (simple)
M_{12} = hazardousness (complex)
M_{21} = effectiveness (simple)
M_{22} = effectiveness (complex)

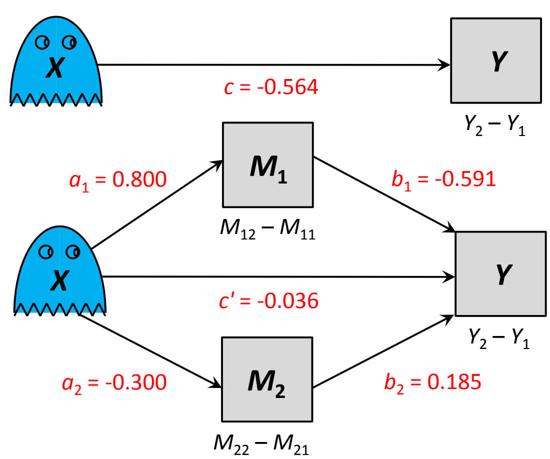


$$Y_2 - Y_1 = c + e_1$$
$$M_{12} - M_{11} = a_1 + e_2$$
$$M_{22} - M_{21} = a_2 + e_3$$
$$Y_2 - Y_1 = c' + b_1(M_{12} - M_{11}) + b_2(M_{22} - M_{21}) + b_3(M_{12} + M_{11})^* + b_4(M_{22} + M_{21})^* + e_4$$

* mean centered

A parallel mediation model in path analytic form

Y_1 = willingness to buy (simple)
Y_2 = willingness to buy (complex)
M_{11} = hazardousness (simple)
M_{12} = hazardousness (complex)
M_{21} = effectiveness (simple)
M_{22} = effectiveness (complex)



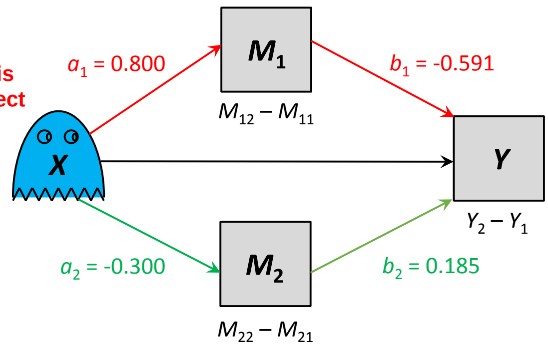
$c = c' + a_1b_1 + a_2b_2$

$$c = \underbrace{-0.036}_{\text{Direct effect}} + \underbrace{(0.800)(-0.591)}_{\text{Indirect effect via } M_1} + \underbrace{(-0.300)(0.185)}_{\text{Indirect effect via } M_2} = \underbrace{-0.036}_{\text{Direct effect}} + \underbrace{-0.473 + -0.055}_{\text{Indirect effects}} = \underbrace{-0.564}_{\text{Total effect}}$$

Statistically comparing indirect effects

Analytical goal: Determine if the indirect effect of name complexity on willingness to buy through hazardousness is different than the indirect effect through effectiveness.

Y_1 = willingness to buy (simple)
Y_2 = willingness to buy (complex)
M_{11} = hazardousness (simple)
M_{12} = hazardousness (complex)
M_{21} = effectiveness (simple)
M_{22} = effectiveness (complex)



Specific indirect effect of name complexity through perceived hazardousness:

$a_1 b_1$= (0.800)(-0.591) = -0.473

Specific indirect effect of name complexity through perceived effectiveness:

$a_2 b_2$= (-0.300)(0.185) = -0.056

We can easily test whether these indirect effects are equal or different using bootstrapping. MEMORE for SPSS and SAS does this test.

MEMORE Output

MEMORE can do all this, including bootstrap confidence intervals for specific indirect effects and their difference.

SPSS:
```
memore y=buy2 buy1/m=hazard2 hazard1 effect2 effect1/contrast=1/samples=10000.
```

SAS:
```
%memore (data=drugname,y=buy2 buy1,m=hazard2 hazard1 effect2 effect1,contrast=1,samples=10000);
```

```
Variables:
Y =    buy2     buy1
M1 =   hazard2  hazard1
M2 =   effect2  effect1

Computed Variables:
Ydiff =           buy2      -      buy1
M1diff =          hazard2   -      hazard1
M2diff =          effect2   -      effect1
M1avg   = (       hazard2   +      hazard1   )      /2      Centered
M2avg   = (       effect2   +      effect1   )      /2      Centered

Sample Size:
  22

********************************************************************************
Outcome: Ydiff =  buy2      -      buy1

Model
         Effect        SE         t        df         p      LLCI      ULCI
'X'      -.5636     .1932   -2.9168   21.0000     .0082    -.9655    -.1618

********************************************************************************
Outcome: M1diff = hazard2   -      hazard1

Model
         Effect        SE         t        df         p      LLCI      ULCI
'X'       .8000     .2579    3.1024   21.0000     .0054     .2637    1.3363

********************************************************************************
```

c path →

a_1 path →

MEMORE Output

****

Outcome: M2diff = effect2 - effect1

Model

		Effect	SE	t	df	p	LLCI	ULCI
a_2 path →	'X'	-.3000	.1798	-1.6683	21.0000	.1101	-.6740	-.0740

****

Outcome: Ydiff = buy2 - buy1

Model Summary

R	R-sq	MSE	F	df1	df2	p
.8212	.6744	.3304	8.8040	4.0000	17.0000	.0005

Model

		coeff	SE	t	df	p	LLCI	ULCI
c' path →	'X'	-.0357	.1517	-.2352	17.0000	.8169	-.3557	.2844
b_1 path →	M1diff	-.5905	.1165	-5.0684	17.0000	.0001	-.8364	-.3447
b_2 path →	M2diff	.1851	.1596	1.1599	17.0000	.2621	-.1516	.5218
	M1avg	-.2898	.1738	-1.6679	17.0000	.1137	-.6564	.0768
	M2avg	-.2361	.1625	-1.4528	17.0000	.1645	-.5791	.1068

****

2016 SPSP ANNUAL CONVENTION #SPSP2016

MEMORE Output

********************** TOTAL, DIRECT, AND INDIRECT EFFECTS **********************

Total effect of X on Y

	Effect	SE	t	df	p	LLCI	ULCI
c path →	-.5636	.1932	-2.9168	21.0000	.0082	-.9655	-.1618

Direct effect of X on Y

	Effect	SE	t	df	p	LLCI	ULCI
c' path →	-.0357	.1517	-.2352	17.0000	.8169	-.3557	.2844

Indirect Effect of X on Y through M

		Effect	BootSE	BootLLCI	BootULCI
a_1b_1 →	Ind1	-.4724	.1469	-.7445	-.1644
a_2b_2 →	Ind2	-.0555	.0964	-.2177	.1943
	Total	-.5280	.1411	-.7695	-.2173

Point estimates and 95% bootstrap confidence intervals for the specific indirect effects. These are consistent with a claim of mediation by **hazardousness** but <u>not</u> **effectiveness**.

Indirect Key

Ind1	X	->	M1diff	->	Ydiff
Ind2	X	->	M2diff	->	Ydiff

Pairwise Contrasts Between Specific Indirect Effects

	Effect	BootSE	BootLLCI	BootULCI
(C1)	-.4169	.2045	-.8744	-.0178

$a_1b_1 - a_2b_2 = -0.472 - -0.055 = -0.417$

Point estimate and 95% bootstrap confidence interval for the difference between the two specific indirect effects. They are statistically different.

Contrast Key:
(C1) Ind1 - Ind2

************************* ANALYSIS NOTES AND WARNINGS

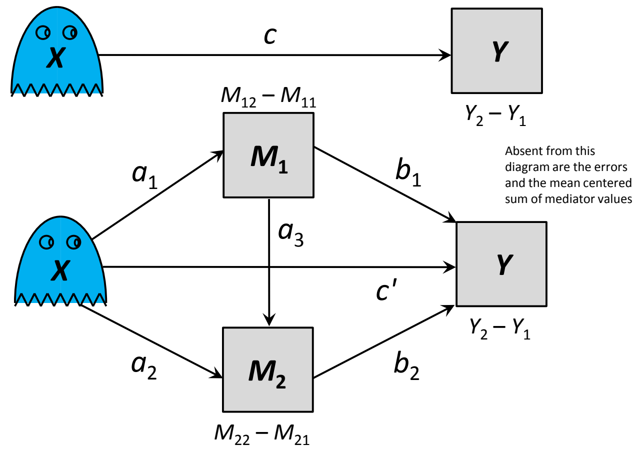
Bootstrap confidence interval method used: Percentile bootstrap.
Number of bootstrap samples for bootstrap confidence intervals: 10000

2016 SPSP ANNUAL CONVENTION #SPSP2016

A serial mediation model in path analytic form

Y_1 = willingness to buy (simple)
Y_2 = willingness to buy (complex)
M_{11} = hazardousness (simple)
M_{12} = hazardousness (complex)
M_{21} = effectiveness (simple)
M_{22} = effectiveness (complex)

$c = c' + a_1 b_1 + a_2 b_2 + a_1 a_3 b_2$



$$Y_2 - Y_1 = c + e_1$$

$$M_{12} - M_{11} = a_1 + e_2$$

$$M_{22} - M_{21} = a_2 + a_3(M_{12} - M_{11}) + b_5(M_{12} + M_{11})^* + e_3$$

* mean centered

$$Y_2 - Y_1 = c' + b_1(M_{12} - M_{11}) + b_2(M_{22} - M_{21}) + b_3(M_{12} + M_{11})^* + b_4(M_{22} + M_{21})^* + e_4$$

MEMORE Output

MEMORE can do all this, including bootstrap confidence intervals for specific indirect effects and their difference.

SPSS:
```
memore y=buy2 buy1/m=hazard2 hazard1 effect2 effect1/contrast=1/serial=1/samples=10000.
```

SAS:
```
%memore (data=drugname,y=buy2 buy1,m=hazard2 hazard1 effect2 effect1,contrast=1,serial=1,samples=10000);
```

```
Variables:
Y =   buy2      buy1
M1 =  hazard2   hazard1
M2 =  effect2   effect1

Computed Variables:
Ydiff =            buy2      -      buy1
M1diff =           hazard2   -      hazard1
M2diff =           effect2   -      effect1
M1avg   = (        hazard2   +      hazard1  )      /2      Centered
M2avg   = (        effect2   +      effect1  )      /2      Centered

Sample Size:
  22

**************************************************************************
Outcome: Ydiff =  buy2      -      buy1

Model
          Effect        SE         t        df          p      LLCI      ULCI
'X'       -.5636     .1932   -2.9168   21.0000      .0082    -.9655    -.1618

**************************************************************************
Outcome: M1diff = hazard2   -      hazard1

Model
          Effect        SE         t        df          p      LLCI      ULCI
'X'        .8000     .2579    3.1024   21.0000      .0054     .2637    1.3363

**************************************************************************
```

c path →

a_1 path →

MEMORE Output

```
*******************************************************************************
Outcome: M2diff = effect2    -     effect1

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .3308      .1094      .7003     1.1675     2.0000    19.0000      .3325
```

Model

		coeff	SE	t	df	p	LLCI	ULCI
a_2 path	'X'	-.1224	.2179	-.5618	19.0000	.5808	-.5785	.3337
a_3 path	M1diff	-.2220	.1563	-1.4200	19.0000	.1718	-.5493	.1052
	M1avg	.0411	.2326	.1766	19.0000	.8617	-.4457	.5278

```
*******************************************************************************
Outcome: Ydiff =  buy2    -     buy1

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .8212      .6744      .3304     8.8040     4.0000    17.0000      .0005
```

Model

		coeff	SE	t	df	p	LLCI	ULCI
c' path	'X'	-.0357	.1517	-.2352	17.0000	.8169	-.3557	.2844
b_1 path	M1diff	-.5905	.1165	-5.0684	17.0000	.0001	-.8364	-.3447
b_2 path	M2diff	.1851	.1596	1.1599	17.0000	.2621	-.1516	.5218
	M1avg	-.2898	.1738	-1.6679	17.0000	.1137	-.6564	.0768
	M2avg	-.2361	.1625	-1.4528	17.0000	.1645	-.5791	.1068

2016 SPSP ANNUAL CONVENTION #SPSP2016

MEMORE Output

********************** TOTAL, DIRECT, AND INDIRECT EFFECTS **********************

Total effect of X on Y

	Effect	SE	t	df	p	LLCI	ULCI
c path	-.5636	.1932	-2.9168	21.0000	.0082	-.9655	-.1618

Direct effect of X on Y

	Effect	SE	t	df	p	LLCI	ULCI
c' path	-.0357	.1517	-.2352	17.0000	.8169	-.3557	.2844

Indirect Effect of X on Y through M

		Effect	BootSE	BootLLCI	BootULCI
a_1b_1	Ind1	-.4724	.1469	-.7445	-.1644
a_2b_2	Ind2	-.0227	.0611	-.1531	.1085
$a_2a_3b_2$	Ind3	-.0329	.0912	-.2401	.1499
	Total	-.5280	.1411	-.7695	-.2173

Point estimates and 95% bootstrap confidence intervals for the specific indirect effects. These results are consistent with a claim of mediation by **hazardousness** alone but <u>not</u> **effectiveness** or hazardousness and effectiveness in serial.

```
Indirect Key
Ind1  X      ->      M1diff   ->      Ydiff
Ind2  X      ->      M2diff   ->      Ydiff
Ind3  X      ->      M1diff   ->      M2diff   ->      Ydiff
```

Pairwise Contrasts Between Specific Indirect Effects

	Effect	BootSE	BootLLCI	BootULCI
(C1)	-.4498	.1595	-.7649	-.1419
(C2)	-.4396	.2033	-.8409	-.0256
(C3)	.0102	.1209	-.2234	.2914

$a_1b_1 - a_2b_2$ = -0.473 – -0.023 = -0.450
$a_1b_1 - a_3b_3$ = -0.473 – -0.033 = -0.440
$a_2b_2 - a_3b_3$ = -0.023 – -0.033 = 0.010

```
Contrast Key:
(C1)  Ind1     -     Ind2
(C2)  Ind1     -     Ind3
(C3)  Ind2     -     Ind3
```

Point estimates and 95% bootstrap confidence intervals for the difference between pairs of specific indirect effects.

************************* ANALYSIS NOTES AND WARNINGS *************************

2016 SPSP ANNUAL CONVENTION #SPSP2016

Summary

- Framing Judd, Kenny and McClelland in a path-analytic framework allows for:
  - Focusing inference about mediation on the indirect effect
  - Use of modern inferential methods for the indirect effect (e.g., bootstrapping)
  - Easy generalization to parallel and serial multiple mediation models.
- PROCESS or MEMORE can be used to easily estimate simple mediation models, including a variety of options for inference.
- MEMORE can be used to estimate parallel and serial multiple mediation models, including many options for inference.
  - Researchers can compare theories about processes, by testing if indirect effects through proposed mediators differ significantly.
- Easy to extend to conditional process models, where the indirect effect is moderated by some other variable .
  - Test for order effects, mixed designs (between and within-subject factors), moderation by individual differences.

Thank you!

Thank you to Simone Dohle and Michael Siegrist for allowing us to use their data! Thanks to the National Science Foundation Graduate Research Fellowship and The Ohio State University Distinguished Dean's University Fellowship for supporting Amanda Montoya.

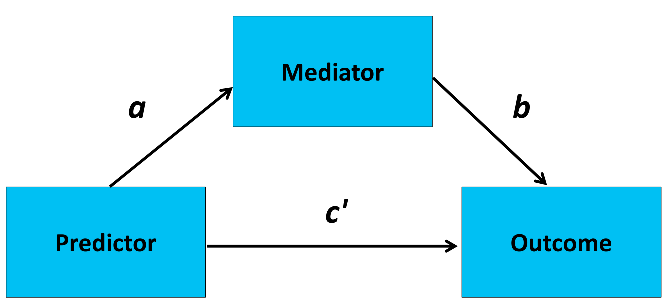
The Mechanisms and Contingencies (MAC) Lab
The Ohio State University

2016 SPSP ANNUAL CONVENTION #SPSP2016

ACCURATE INDIRECT EFFECTS IN MULTILEVEL MEDIATION FOR REPEATED MEASURES DATA

Amanda Sharples and Elizabeth Page-Gould
University of Toronto

Mediation

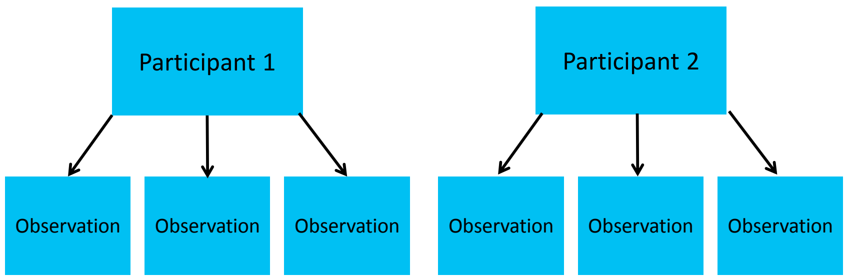


Indirect effect = $a \times b$
Total effect = Indirect effect + c'

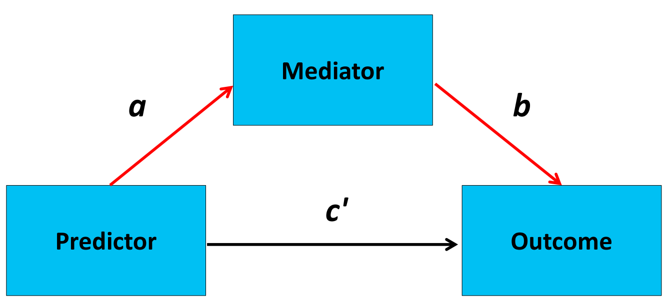
2016 SPSP ANNUAL CONVENTION #SPSP2016

Multilevel Models

Nested (Repeated Measures) Data



Multilevel Mediation

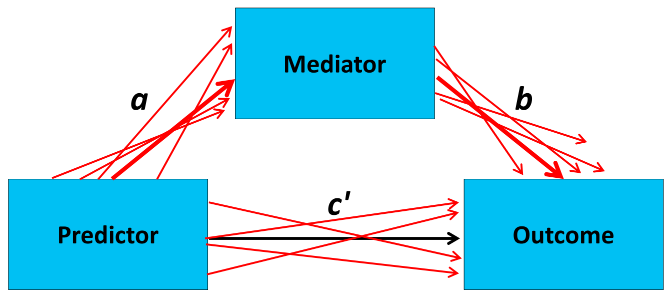


Indirect effect = $a \times b$

Total effect = Indirect effect + c'

The Wrong Way to Do Multilevel Mediation

USE FIXED SLOPES TO CALCULATE INDIRECT EFFECT



Indirect effect = $a \times b$
Total effect = Indirect effect + c'

Why is this Bad?

- The indirect effect is biased.
    - So the total effect is biased too.
- They are biased by how much the random slopes *a* and *b* covary.

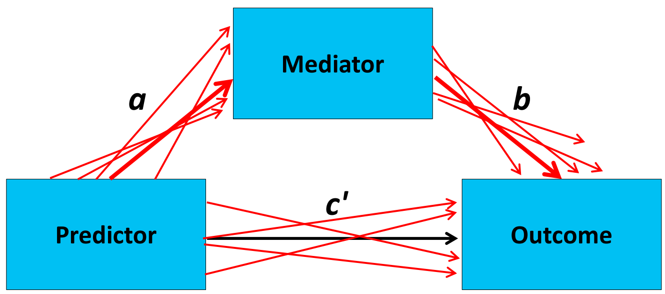
Bauer, Preacher, & Gil (2006); Kenny, Korchmaros, and Bolger (2003)

Bias = $\mathrm{COV}(a_i, b_i) = \sigma_{ab}$
Real indirect effect = $(a \times b) + \mathrm{COV}(a_i, b_i)$
Real total effect = $(a \times b) + \mathrm{COV}(a_i, b_i) + c'$

2016 SPSP ANNUAL CONVENTION #SPSP2016

The Right Way to Do Multilevel Mediation

TAKE RANDOM SLOPES INTO ACCOUNT

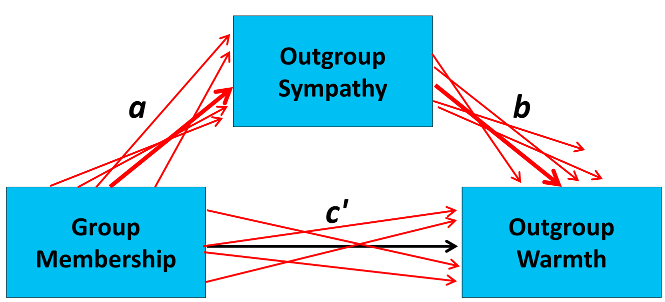


Indirect effect = Mean($a_i \times b_i$)

Total effect = Mean(Indirect effect$_i$ + c'_i)

The Right Way to Do Multilevel Mediation

TAKE RANDOM SLOPES INTO ACCOUNT



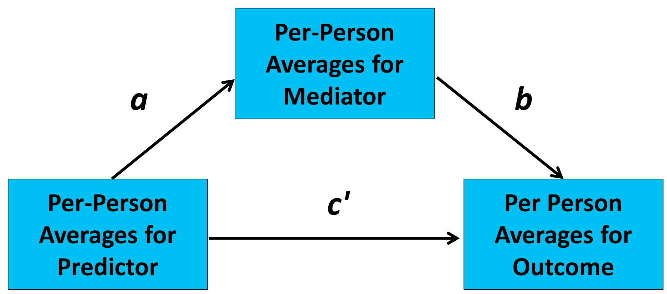
Indirect effect = Mean($a_i \times b_i$)

Total effect = Mean(Indirect effect$_i$ + c'_i)

2016 SPSP ANNUAL CONVENTION
#SPSP2016

An OK Way to Do Multilevel Mediation

USE AGGREGATE REPEATED MEASURES FOR EACH PARTICIPANT

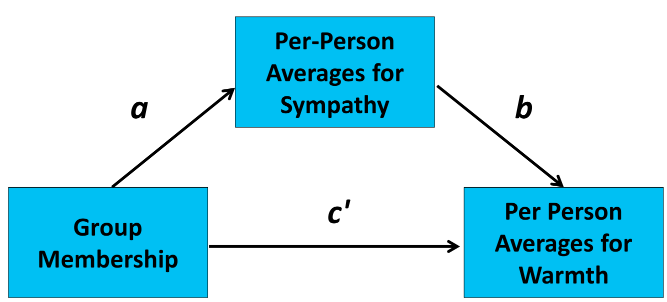


(Unbiased) Indirect effect = $a \times b$
(Unbiased) Total effect = Indirect effect + c'

2016 SPSP ANNUAL CONVENTION #SPSP2016

An OK Way to Do Multilevel Mediation

USE AGGREGATE REPEATED MEASURES FOR EACH PARTICIPANT



(Unbiased) Indirect effect = $a \times b$
(Unbiased) Total effect = Indirect effect + c'

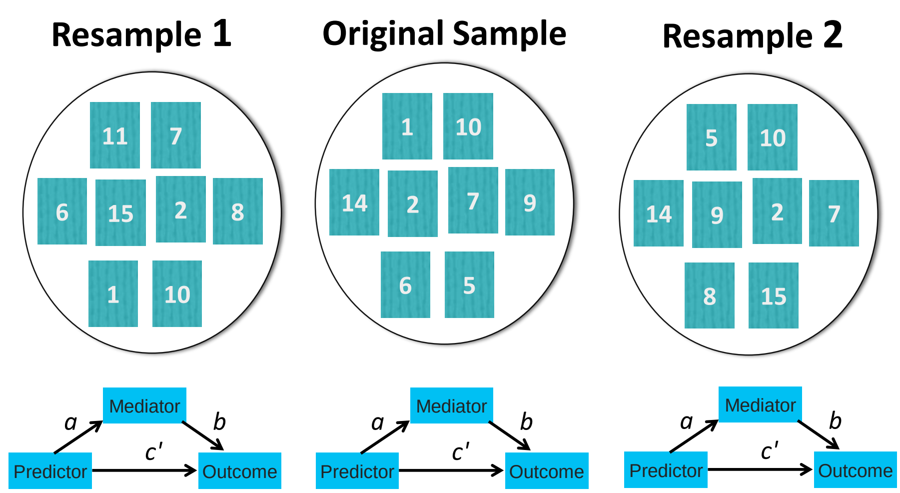
2016 SPSP ANNUAL CONVENTION #SPSP2016

How do we determine the robustness of our effects?

- There have been approaches put forward, but…

- Bootstrapping is ideal because
  - It does not require the assumption that the random effects are normally distributed.
  - It is already ubiquitous in social psychology (especially in mediation analysis)

Bootstrapping for confidence intervals



2016 SPSP ANNUAL CONVENTION #SPSP2016

Goals of Current Demonstration

- Demonstrate how you can calculate unbiased indirect and total effects in multilevel mediation models.
- Demonstrate how you can use a bootstrapping approach to estimate confidence intervals for your effects.

Research Questions

- Will people rate their target in-group more warmly than target outgroups?
- Can this be explained by greater sympathy toward the target in-group (i.e., an indirect effect).

Method: Sample

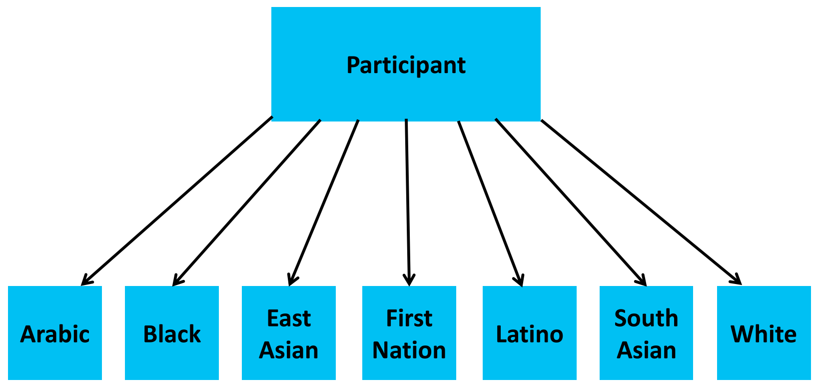
- N = 340 (community members)
- 62% female, 38% male
- Age range: 16-75
- Ethnicity: 33% White, 28% East Asian, 28% South Asian, 5% Black, 3% Arab, 2% Latino

Method: Questionnaire

- Demographic information (e.g., ethnicity).
- Sympathy (0 = not at all sympathetic to 10 = very sympathetic) toward 7 target ethnic groups.
- Warmth (0 = cold to 10 = warm) toward 7 target ethnic groups.

Arabic	Black	East Asian	First Nation	Latino	South Asian	White

Analytic Approach



Method: Questionnaire

Bootstrap Analysis in R:

- Created a function "indirect.mlm"
  - Runs the relevant multilevel models in each resample
  - Multiplies together the <u>random</u> *a* and *b* slopes and takes the mean of these products
- Use the "boot" package to do the multilevel mediation

Analytic Approach

Between-Person Effects:

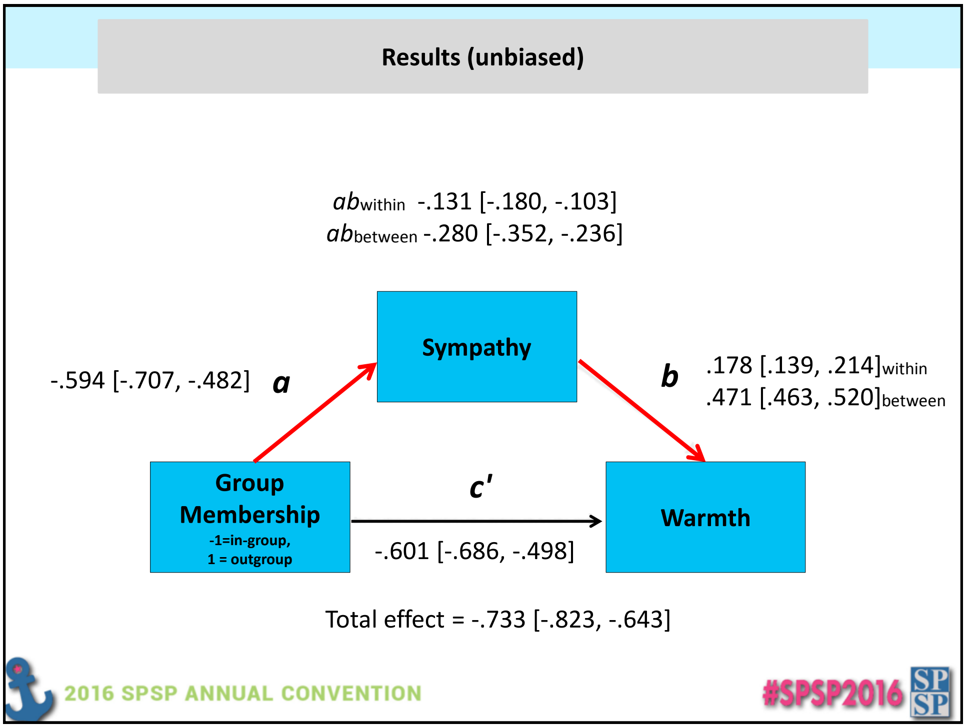
- Indirect effect = $a \times b$
- Total effect = Indirect effect + c'

Within-Person Effects:

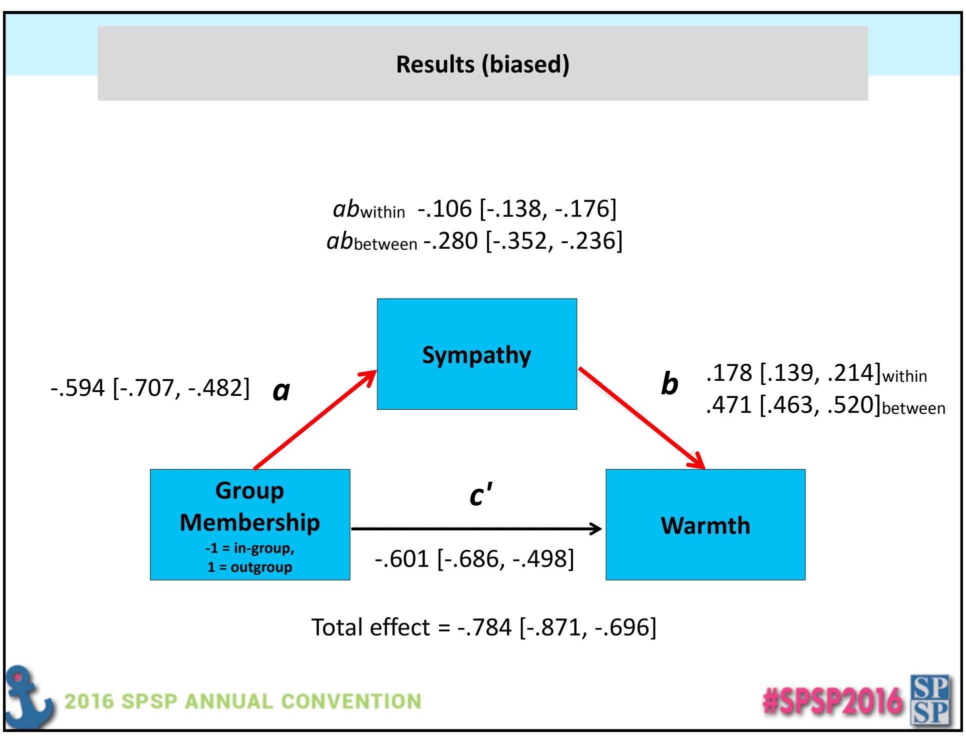
- Unbiased Indirect effect = Mean($a_i \times b_i$)
- Unbiased Total effect = Mean(Indirect effect$_i$ + c'_i)

Analytic Approach

```
boot(data=data.set, R=1000,
strata=ID,
statistic=indirect.mlm,
y="warmth", x="target",
m="sympathy", group.id="ID",
between.m=T,
uncentered.x=F)
```



Results (unbiased)
ab within -.131 [-.180, -.103]
ab between -.280 [-.352, -.236]
Sympathy
-.594 [-.707, -.482] a
b .178 [.139, .214] within
.471 [.463, .520] between
Group Membership
-1=in-group,
1 = outgroup
c'
Warmth
-.601 [-.686, -.498]
Total effect = -.733 [-.823, -.643]
2016 SPSP ANNUAL CONVENTION
#SPSP2016



Results (biased)
ab within -.106 [-.138, -.176]
ab between -.280 [-.352, -.236]
Sympathy
-.594 [-.707, -.482] a
b .178 [.139, .214] within
.471 [.463, .520] between
Group Membership
-1 = in-group,
1 = outgroup
c'
Warmth
-.601 [-.686, -.498]
Total effect = -.784 [-.871, -.696]
2016 SPSP ANNUAL CONVENTION
#SPSP2016

Results

Bias in indirect effect:

Biased: ab_{within} = -.106 [-.138, -.076]
Unbiased: ab_{within} = -.131 [-.180, -.103]

Difference = .025 [.015, .058] = σ_{ab}

- Difference between biased and unbiased effects is equal to covariance between random slopes for paths *a* and *b*.

Bauer et al. (2006)

Bauer et al. (2006)

Results

Bias in total effect:

Biased: c = -.784 [-.871, -.696]
Unbiased: c = -.733 [-.823, -.643]

Difference = -.052 [-.086, -.020]

- Difference between biased and unbiased total effect is equal to

$$ab_{unbiased} - ab_{biased} + \sigma_{ab}$$

Bauer et al. (2006)

Bauer et al. (2006)

Discussion

- Download R script to run this analysis
  - www.page-gould.com/r/indirectmlm
- Currently, SPSS doesn't allow you to save random slopes in its MIXED procedure
  - You can't do this analysis in SPSS right now.
  - IBM says this is planned for future release.
- Good news!
  - We are creating a web application for non-R users.

Take Home Message

- Proof of concept
  - You can bootstrap indirect effects in multilevel mediation analysis.

www.page-gould.com/r/indirectmlm

Thank you!

Co-author:

Elizabeth Page-Gould

- Lab and Research Assistants:
  - Social Psychophysiology and Quantitative Methods Lab (SPRQL)
- Funding Sources:
  - Awarded to Sharples:
    - Ontario Graduate Scholarship

Awarded to Page-Gould:

- Canada Research Chairs
- Canada Foundation for Innovation
- Connaught Fund New Researcher Award
- Ontario Ministry of Research & Innovation
- Social Sciences and Humanities Research Council (SSHRC) Insight Grants

2016 SPSP ANNUAL CONVENTION

#SPSP2016