
# Using MEMORE:
## Mediation in Two-Instance Repeated Measures Designs

Amanda Kay Montoya
Graduate Student
Ohio State University

The Ohio State University
Department of Psychology
Social Psychology Methodology Workshops
September 8th, 2016

1

## Introductions

Please go around the room and state:

1) Your Name
2) Your Program and Year
3) What experience do you have with mediation?
4) What experience do you have with repeated-measures data?

While we're doing this please go to https://github.com/akmontoya/OSUSocial and download the folder and open SPSS.

2

## Workshop Procedures

- One Hour Workshop on Mediation in Two-Instance Within-Participant Designs

- Assuming some familiarity with:
  - Regression
  - Mediation
  - SPSS (or SAS)

- Download with data files and syntax for SPSS [INSERT LINK]
  - MEMORE version 1.1
    - Available at *http://www.akmontoya.com/spss-and-sas-macros*

Instruction Procedures

  - Combination of theory and practice
  - Follow along with the analysis as we go

```
SPSS Code
```

    - Use syntax!
    - **Ask questions** about concepts or anything that is confusing

  - Make friends with your neighbors, so if you have troubles as you go through you can work together



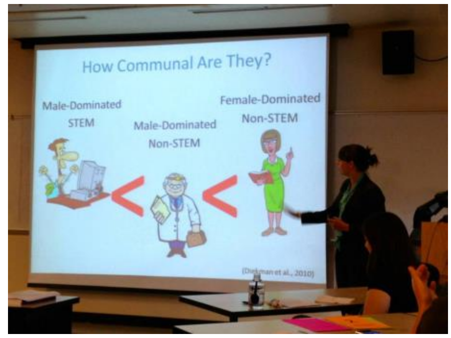
## Workshop Topics

- Assumed familiarity with linear regression and some familiarity with mediation
- Between Subjects Mediation ← Slow me down here if needed!
  - Path analytic approach
  - Estimation
  - Interpretation
  - Inference
- Repeated Measures Data
- Two-Condition Within Subjects Mediation
  - Setting up the regression equations
  - Path Analysis
  - Estimation of Indirect Effects
  - Inference
  - Using MEMORE
  - Common Questions
  - Reporting (Writing and Figures)
- Other Types of Repeated Measures Mediation
  - Multilevel (1 – 1 – 1 , 1 – 2 – 2 etc)
  - Longitudinal
  - Multilevel SEM

## Running Example

Montoya, A. K. (2013) Increasing Interest in Computer Science thought Group Work: A Goal Congruity Approach (Undergraduate Honors Thesis).

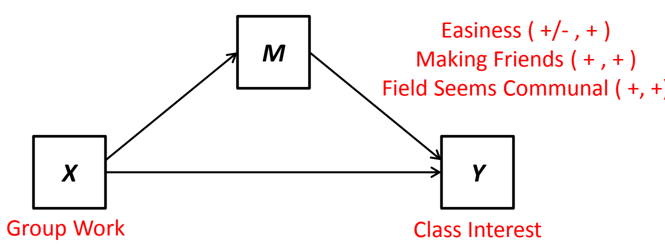
**Can group work in computer science classes increase women's interest by increasing their perception that computer science is communal?**




**Between-Subjects Version (CASC_BS.sps) :**
Female participants (N = 107) read one of two syllabi for a computer science class. One of the syllabi reported the class would have group projects throughout, and the other syllabi stated that individual project would be scheduled throughout.

**Measured Variables:**
- Interest in the class (aggregate over 4 items)
- Perceptions that computer science is a communal field (aggregate over 5 items)

## Mediation




A simple mediation model connects an assumed causal variable (*X*) to an assumed outcome variable (*Y*), through some mechanism (*M*). The effect of *X* on *Y* can move completely through *M* or only partially through *M*.

*M* is frequently referred to as a *mediator* or *intermediary variable*.

Many different kind of variables may act as mediators. Emotional variables, situational, individual level variables, cognitive variables, environmental variables, etc.

Mediation can be found throughout the psychology literature

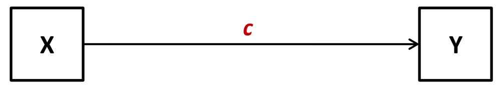
A quick example: Name some possible mediators!

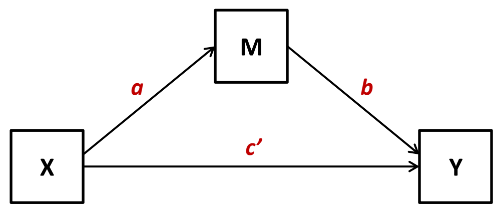
## Mediation: Path Analysis

Consider *a*, *b*, *c*, and *c′* to be measures of the causal effect of the variables in the mediation model.

These could be measured using regression coefficients from OLS or path estimates in a structural equation model using maximum likelihood estimation.

$$Y_i = i_{Y^*} + cX_i + e_{Y_i^*}$$
$$M_i = i_M + aX_i + e_{M_i}$$
$$Y_i = i_Y + c'X_i + bM_i + e_{Y_i}$$




Direct effect of *X* on *Y* (not through *M*) = *c′*

Indirect effect of *X* on *Y* (through *M*) = *a* × *b*

Total effect = direct effect + indirect effect
*c* = *c′* + *a* × *b*

Indirect effect = total effect - direct effect
*a* × *b* = *c* - *c′*

7

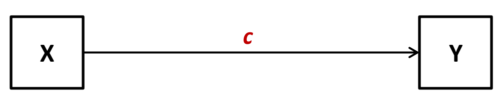
## Interpreting the Coefficients

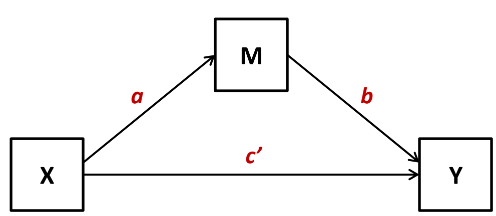
**Total Effect (*c*):** The effect of our presumed cause (*X*) on our outcome (*Y*), without controlling for any other variables.

***a*-path:** The effect of our presumed cause (*X*) on our mediator *(M)*.

***b*-path:** The effect of our mediator (*M*) on the outcome (*Y*) while controlling for *X*. (i.e. predicted difference in *Y* for two people with the same score on X but who differ on *M* by one unit).

**Direct effect (*c′*):** The effect of our presumed cause (*X*) on *Y* while controlling for *M*. (i.e. predicted difference in *Y* for two people who differ by one unit on *X* but with the same score on *M*)



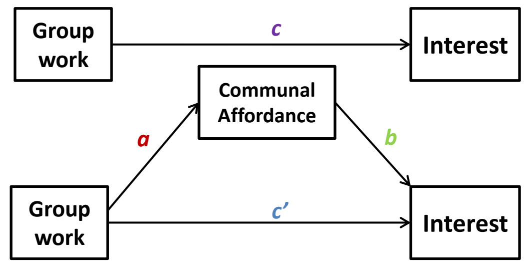




**Indirect Effect (*ab*):** Product of effect of *X* on *M*, and effect of *M* on *Y* controlling for *X*. The effect of *X* on *Y* through *M*.

8

## Estimation with CASC BS Data

Consider the question **Is the effect of group work on interest mediated by communal affordance?**




The *c*-path can be estimated in a sample using the regression equation below.

$$Y_i = i_{Y^*} + cX_i + e_{Y_i^*}$$

```
regression /dep = interest /method = enter gw.
```

**Coefficients[a]**

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. |
|---|---|---|---|---|---|---|
| 1 | (Constant) | 2.239 | .439 | | 5.098 | .000 |
| | gw | .462 | .285 | .156 | 1.621 | .108 |

a. Dependent Variable: interest

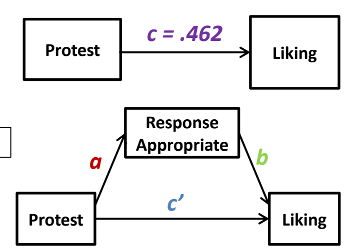
Overall women were .462 units more interested in the class with group work.

$c$ = .462



## Estimation with Protest Data

**Is the effect of protesting on liking mediated by perceptions of response appropriateness?**




$$M_i = i_M + aX_i + e_{M_i}$$

```
regression /dep = comm /method = enter gw.
```

**Coefficients[a]**

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. |
|---|---|---|---|---|---|---|
| 1 | (Constant) | 2.933 | .364 | | 8.067 | .000 |
| | gw | .488 | .237 | .198 | 2.060 | .042 |

a. Dependent Variable: comm

$a$ = .488

Women saw computer science as .488 units more communal after reading a syllabus with group work.

$$Y_i = i_Y + c'X_i + bM_i + e_{Y_i}$$

```
regression /dep = interest /method = enter Cond comm.
```

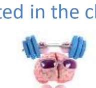
**Coefficients[a]**

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. |
|---|---|---|---|---|---|---|
| 1 | (Constant) | .746 | .515 | | 1.448 | .151 |
| | gw | .218 | .268 | .073 | .812 | .419 |
| | comm | .508 | .109 | .421 | 4.663 | .000 |

a. Dependent Variable: interest

Controlling for communal affordance, women in the group work condition were .218 units more interested in the class.

$c'$ = .218



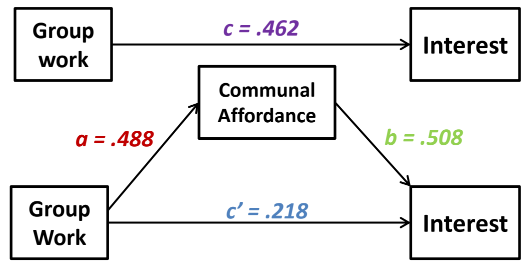
$b$ = .508

Interpretation?

## Estimation with Protest Data

**Is the effect of protesting on liking mediated by perceptions of response appropriateness?**

On average, women were .46 units more interested in the class with group work ($p$ = .108). Similarly, computer science was perceived as .49 units more communal after reading a syllabus with group work ($p$ = .042). Controlling for condition, a one unit increase in communal affordance resulted in a .508 unit increase in interest ($p$ < .001). Controlling for communal affordance, group work did not predict additional interest ($c'$ = .22, $p$ = .42).




**But what about the indirect effect?**

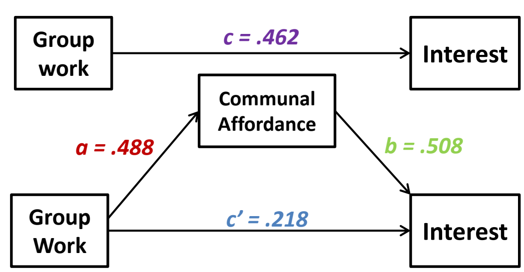


## Estimation with Protest Data

**Indirect Effect**

$a \times b$ = .488 × .508 = .249

Group work increased interest by .249 units indirectly through communal affordance. Where group work increased perceptions of communal affordance by .488 units, and a one unit increase in communal affordance resulted in a .508 unit increase in interest.




**Direct Effect**

$c'$ = .218

Group work increased interest by .218 units directly (not through communal affordance).

**Total Effect**

$c$ = .462

Group work increased interest by .462 units in total.

*Inference for the direct and total effects can be drawn from the regression results because these are based on a single regression parameter.*

$p$ = . 419

$p$ = .108

## Inference about the Indirect Effect

- How to make proper inference about the indirect effect may be the most active area of research in mediation analysis
- Some methods you may have heard of
  - Causal Steps / Baron and Kenny Method / Baron and Kenny Steps
  - Test of Joint Significance
  - Sobel Test / Multivariate Delta Method
  - Monte Carlo Confidence Intervals
  - Distribution of the Product Method
  - Bootstrap Confidence Intervals
    - Percentile Bootstrap
    - Bias-Corrected Bootstrap
    - Bias Corrected and Accelerated Bootstrap

  I'm just going to talk about bootstrapping

- Why is this so hard?

  The product of two normal distributions is not necessarily normal. There are many instances where the indirect effect could be zero (either $a$ or $b$ could be zero, or both could be zero).



## Causal Steps Method

**Method**

1. Test if there is a significant total effect ($c \neq 0$).
2. Test if there is a significant effect of $X$ on $M$ ($a \neq 0$).
3. Test if there is a significant effect of $M$ on $Y$ controlling for $X$ ($b \neq 0$).
4. If all three steps are confirmed, test for partial vs. complete mediation.
   1. If $X$ still has an effect on $Y$ controlling for $M$ ($c' \neq 0$), this is partial mediation
   2. If $X$ does not have a significant effect on $Y$ controlling for $M$, complete mediation

**Appeal**

- Easy to do, just need regression
- Intuitive

**What's wrong with it?**

- No estimate of the indirect effect
- No quantification of uncertainty about conclusion
  - *p*-value
  - Confidence Interval
- Requirement that the total effect is significant before looking for indirect effect
- Multiple testing problem
- Issues with *complete* and *partial* mediation

## Bootstrap Confidence Intervals
## (Percentile)

Empirically estimate sampling distribution of the indirect effect. From this distribution compute confidence intervals which can be used for estimation and hypothesis testing.

**Bootstrapping the Indirect Effect**

1. Randomly sample *n* cases from your dataset with replacement.
2. Estimate the indirect effect using resampled dataset, call this $ab^{(1)}$
3. Repeat steps 1 and 2 a total of *K* times where *K* is many (10,000 recommended), each time calculated $ab^{(k)}$.
4. The sampling distribution of the $ab^{(i)}$'s can be used as an estimate of the sampling distribution of the indirect effect.
5. For a 95% confidence interval the lower and upper bounds will be the $2.5^{th}$ and $97.5^{th}$ percentiles of the *K* estimates of the indirect effect.

**Appeal**

- No assumptions about the sampling distribution of the indirect effect
- Provides point estimate of indirect effect
- Can calculate confidence intervals
- Most simulation work suggests this is a very good method which balances Type I Error and Power

**What's wrong with it?**

- Most software does not have this functionality built in
- Requires original data

15

## Bootstrap Confidence Intervals

Original Data

| X | M | Y |
|---|---|---|
| -0.35 | -0.58 | 0.25 |
| 0.31 | -0.50 | 1.89 |
| -0.19 | 2.61 | 2.08 |
| -1.30 | -1.49 | -0.54 |
| 0.59 | 1.14 | 1.74 |
| -0.29 | -0.29 | 1.04 |
| 1.80 | 0.08 | 1.23 |
| -0.01 | 1.20 | 1.30 |
| 0.30 | 1.35 | 1.31 |
| -0.98 | 0.90 | -0.76 |

Bootstrap Sample

| X | M | Y |
|---|---|---|
| -0.35 | -0.58 | 0.25 |
| -0.19 | 2.61 | 2.08 |
| 0.30 | 1.35 | 1.31 |
| 0.59 | 1.14 | 1.74 |
| 0.31 | -0.50 | 1.89 |
| -0.01 | 1.20 | 1.30 |
| 0.30 | 1.35 | 1.31 |
| 0.31 | -0.50 | 1.89 |
| 0.30 | 1.35 | 1.31 |
| -0.01 | 1.20 | 1.30 |

Original data: $a$ = .2931 $b$ = .3099

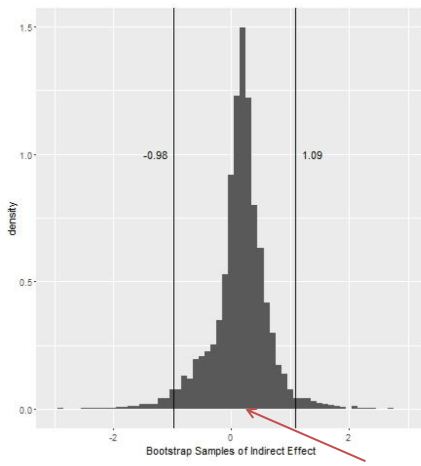
$ab$ = .0908

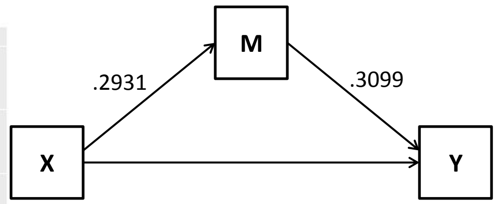
Bootstrap sample: $a$ = -.1035 $b$ = .1495

$ab$ = -.0155

16

## Bootstrap Confidence Intervals




*ab* = .0908

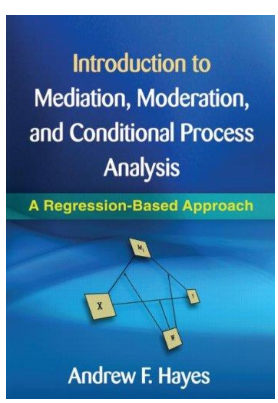
Zero is contained in the confidence interval so we **cannot** claim an indirect effect different from zero with 95% confidence. This is the same as failing to reject the null hypothesis at α = .05.

*ab* = .0908



## PROCESS

PROCESS is a macro available for SPSS and SAS written by Andrew F. Hayes, documented in *Mediation, Moderation, and Conditional Process Analysis*, and available for free online at *processmacro.org*




**Published in May 2013 and available through The Guilford Press, Amazon.com, and elsewhere.**

- PROCESS integrates a variety of macros previously developed by Hayes: SOBEL, INDIRECT, MODMED, MODPROBE, MED3C. If you are using any of these now, switch to PROCESS.
- Current version is 2.16
- PROCESS can assess a variety of models. Find the model you are interested in in the templates file, then use that model number.
- Appendix A of IMMCPA provides complete documentation of options in PROCESS and how to use them.

# Using PROCESS for CASC BS data

```
PROCESS vars = gw interest comm /x = gw /m = comm /y = interest / model =
    4 /percent = 1 /total = 1.
```

```
************** PROCESS Procedure for SPSS Release 2.16.1 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2013). www.guilford.com/p/hayes3

**************************************************************************
Model = 4
    Y = interest
    X = gw
    M = comm

Sample size
        106

**************************************************************************
Outcome: comm

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .1980      .0392     1.4720     4.2428     1.0000   104.0000      .0419

Model
              coeff         se          t          p       LLCI       ULCI
constant     2.9330      .3636     8.0672      .0000     2.2121     3.6540
gw            .4876      .2367     2.0598      .0419      .0182      .9571
```

Outcome variable

Predictors

First part of output repeats what you told PROCESS to do. Always double check that this is correct!

Each chunk of the output like this is one of the regression models involved in the mediation analysis This is the model of *M* from *X*, therefore this is the model which produces the estimate of *a*

19

# Using PROCESS for CASC BS data

```
PROCESS vars = gw interest comm /x = gw /m = comm /y = interest / model = 4 /percent =
    1 /total = 1.
```

```
**************************************************************************
Outcome: interest

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .4411      .1945     1.8165    12.4377     2.0000   103.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant      .7459      .5150     1.4484      .1506     -.2755     1.7672
comm          .5079      .1089     4.6628      .0000      .2919      .7240
gw            .2178      .2683      .8119      .4187     -.3143      .7499

************************** TOTAL EFFECT MODEL ****************************
Outcome: interest

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .1565      .0245     2.1788     2.6122     1.0000   104.0000      .1091

Model
              coeff         se          t          p       LLCI       ULCI
constant     2.2356      .4423     5.0542      .0000     1.3585     3.1128
gw            .4655      .2880     1.6162      .1091     -.1056     1.0367
```

*b* = .508

*c'* = .218

*c* = .466

This is the model predicting *Y* from *X* and *M*, therefore this model gives us an estimate of *b* and *c'*

This is the model predicting *Y* from just *X*. It does not get printed by default but by toggling `total = 1` we can get it out of PROCESS

20

# Using PROCESS for CASC BS data

```
process vars = liking protest respappr /y=liking/x=protest/m=respappr/ total=1
     /percent=1/normal=1/model=4/boot=10000.
```

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

Total effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .4655      .2880     1.6162      .1091     -.1056     1.0367

Direct effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .2178      .2683      .8119      .4187     -.3143      .7499

Indirect effect of X on Y
          Effect    Boot SE   BootLLCI   BootULCI
comm       .2477      .1252      .0213      .5117

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

Number of bootstrap samples for percentile bootstrap confidence intervals:
     5000

Level of confidence for all confidence intervals in output:
    95.00

NOTE: Some cases were deleted due to missing data.  The number of such cases was:
  1

------ END MATRIX -----
```

Indirect effect estimate with a bootstrap confidence interval

Important effects for mediation and inference about these effects

Details about the analysis pop up here

21

# Repeated Measures Data

There are many different kinds of “repeated measures data.” What type of data you have will determine what kind of mediation analysis is appropriate.

Types of *Repeated Measurements:*

- Each person *over time*
- *Nested/Multilevel* data (individuals within schools, cohorts, etc)
- *Dyadic* data (twins, couples, labmates, roomates)
- Each person in a *variety of circumstances*
- and many more…

What is measured repeatedly?

- Specifically in mediation, it’s important to think about how/when/how many times  the variables in your mediation model are measured
- *Multilevel* has a nice system referring to levels (1-1-1 mediation, 1-2-1, mediation etc.
- Is your causal variable measured repeatedly?
- Is your causal variable what differentiates your repeated measurements?

22

## Repeated Measures Data

MEMORE is for two-instance repeated measures mediation analysis, where the causal variable of interest is the factor which differs by repeated measures.

X: varies between repeated measurements
M: measured in each of the two instances
Y: measured in each of the two instances

Examples:
- Participants read two scenarios. Interested in how scenario influences Y through M. Measure M and Y in each scenario.
- Pre-post test: Therapist measures certain symptoms and various outcomes before administering some intervention, and after administering the intervention.
- Researcher interested in if male partners in heterosexual relationships believe fights are less severe because they are less perceptive that small "squabbles" occur. Measure both male and female partners in relationships, self report number of small "squabbles" and severity of last fight.

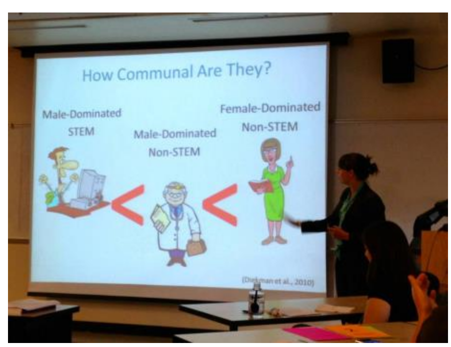
Non-Examples:
- Does calories consumption impact body image through weight gain over time?
- Any instance where nesting is a "nuisance" (e.g. studying schools, but not interested in comparing schools directly).

23

## Running Example

Montoya, A. K. (2013) Increasing Interest in Computer Science thought Group Work: A Goal Congruity Approach (Undergraduate Honors Thesis).

**Can group work in computer science classes increase women's interest by increasing their perception that computer science is communal?**




**Within-Subjects Version (CASC_WS.sps) :**
Female participants (N = 51) read two syllabi for a different computer science classes. One of the syllabi reported the class would have group projects throughout, and the other syllabi stated that individual project would be scheduled throughout.

**Measured Variables:**
- Interest in each the class (aggregate over 4 items)
- Perceptions that the class has a communal environment (aggregate over 5 items)

## 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

Analyses designed to detect mediation and moderation of treatment effects are increasingly prevalent in research in psychology. The mediation question concerns the processes that produce a treatment effect. The moderation question concerns factors that affect the magnitude of that effect. Although analytic procedures have been reasonably well worked out in the case in which the treatment varies between participants, no systematic procedures for examining mediation and moderation have been developed in the case in which the treatment varies within participants. The authors present an analytic approach to these issues using ordinary least squares estimation.

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.

## Within-Subjects Data Example

Montoya, A. K. (2013) Increasing Interest in Computer Science thought Group Work: A Goal Congruity Approach (Undergraduate Thesis).

**Analytic Goal: Can <u>group work</u> in computer science classes increase women's <u>interest</u> by increasing their perception that computer science is <u>communal</u>?**

1. On average, is interest higher in the group work condition?
2. On average, is communal goal affordance higher in the group work condition?
3. Does difference in communal affordance predict a difference in interest?
4. Does the difference in communal goal affordance account for all the difference in interest?

| int_I | int_G | comm_I | comm_G |
|---|---|---|---|
| 1.50 | 4.00 | 1.00 | 6.80 |
| 2.75 | 3.25 | 2.00 | 5.40 |
| 5.75 | 2.50 | 3.20 | 3.60 |
| 3.50 | 5.75 | 1.60 | 5.20 |
| 2.25 | 2.00 | 4.40 | 4.60 |
| 1.50 | 1.75 | 3.00 | 5.00 |
| 2.50 | 4.25 | 4.20 | 4.40 |
| 6.00 | 1.75 | 4.80 | 2.40 |
| 3.00 | 2.00 | 2.60 | 5.80 |
| 4.00 | 5.25 | 1.60 | 5.00 |
| 5.00 | 5.00 | 4.60 | 6.20 |
| 2.00 | 1.75 | 3.80 | 4.20 |
| 1.00 | 1.75 | 2.60 | 3.20 |
| 1.25 | 4.50 | 1.00 | 6.00 |
| 5.75 | 4.50 | 2.60 | 6.00 |
| 3.25 | 4.75 | 3.00 | 6.20 |
| 2.75 | 2.25 | 4.80 | 4.60 |
| 5.50 | 2.00 | 4.00 | 7.00 |
| 1.75 | 5.25 | 1.60 | 5.60 |
| 4.00 | 5.50 | 1.80 | 5.40 |
| 2.25 | 4.00 | 2.20 | 4.80 |
| 4.00 | 6.50 | 2.00 | 6.80 |
| 5.00 | 4.50 | 3.20 | 6.00 |
| 5.00 | 3.75 | 4.00 | 4.80 |
| 4.75 | 5.25 | 1.20 | 6.60 |

# Analysis using Judd et al. (2001)

1. On average, is *Y* higher in one condition than another?

Setup a model of the outcome in each condition:

$$Y_{1i} = c_1 + \epsilon_{Y*1i}$$
$$Y_{2i} = c_2 + \epsilon_{Y*2i}$$

Based on these models, setup a new model where you can directly estimate and conduct inference on what you are interested in (in this case $c_2 - c_1$):

$$Y_{2i} - Y_{1i} = (c_2 - c_1) + (\epsilon_{Y*2i} - \epsilon_{Y*1i}) = c + \epsilon_{Y*i}$$

Use intercept only regression analysis, or a paired sample t-test, or a one sample t-test on the differences to conduct inference on $c_2 - c_1$

With the data: On average, is communal goal affordance higher in the group work condition?

```
T-TEST PAIRS=int_G WITH int_I (PAIRED).
```

**Paired Samples Test**

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference Lower | 95% Confidence Interval of the Difference Upper | | | |
| Pair 1 | int_G - int_I | .37255 | 1.99585 | .27948 | -.18879 | .93389 | 1.333 | 50 | .189 |

# Analysis using Judd et al. (2001)

2. On average, is *M* higher in one condition than another?

Setup a model of the mediator in each condition:

$$M_{1i} = a_1 + \epsilon_{M1i}$$
$$M_{2i} = a_2 + \epsilon_{M2i}$$

Based on these models, setup a new model where you can directly estimate and conduct inference on what you are interested in (in this case $a_2 - a_1$):

$$M_{2i} - M_{1i} = (a_2 - a_1) + (\epsilon_{M2i} - \epsilon_{M1i}) = a + \epsilon_{Mi}$$

Use intercept only regression analysis, or a paired sample t-test, or a one sample t-test on the differences to conduct inference on $a_2 - a_1$

With the data: On average, is communal goal affordance higher in the group work condition?

```
T-TEST PAIRS=comm_G WITH comm_I (PAIRED).
```

**Paired Samples Test**

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference Lower | 95% Confidence Interval of the Difference Upper | | | |
| Pair 1 | comm_G - comm_I | 2.29412 | 1.77870 | .24907 | 1.79385 | 2.79438 | 9.211 | 50 | .000 |

# Analysis using Judd et al. (2001)

3. Does *M* predict *Y* controlling for condition?

Setup a model of the outcome in each condition:

$$Y_{1i} = g_{10} + g_{11}M_{1i} + \epsilon_{Y1i}$$
$$Y_{2i} = g_{20} + g_{21}M_{2i} + \epsilon_{Y2i}$$

Note that there are **two estimates** of the effect of *M* on *Y*. Let's average them to estimate an average effect of *M* on *Y*. Setup a new model where you can directly estimate and conduct inference on what you are interested in (in this case $\frac{1}{2}(g_{11} + g_{21})$):

$$Y_{2i} - Y_{1i} = (g_{20} - g_{10}) + g_{21}M_{2i} - g_{11}M_{1i} + (\epsilon_{Y2i} - \epsilon_{Y1i})$$

Optional board work

$$Y_{2i} - Y_{1i} = (g_{20} - g_{10}) + \underbrace{\frac{g_{21}+g_{11}}{2}}_{b_1}(M_{2i} - M_{1i}) + \underbrace{\frac{(g_{21}-g_{11})}{2}}_{b_2}(M_{2i} + M_{1i}) + (\epsilon_{Y2i} - \epsilon_{Y1i})$$

Next we center the sum term, so the intercept has the interpretation of the predicted difference in *Y* for someone with no difference in *M*'s but is average on *M*'s.

$$Y_{2i} - Y_{1i} = c' + b_1(M_{2i} - M_{1i}) + b_2(M_{2i} + M_{1i} - (\overline{M_2 + M_1}) + (\epsilon_{Y2i} - \epsilon_{Y1i})$$

where $c' = (g_{20} - g_{10} + b_2(\overline{M_2 + M_1}))$



# Analysis using Judd et al. (2001)

3. Does *M* predict *Y* controlling for condition?

With the data: Does communal goal affordance predict interest in the class?

```
compute int_diff = int_G - int_I.
compute comm_diff = comm_G - comm_I.
compute comm_sum = comm_G+comm_I - 8.325490.
EXECUTE.
regression dep = int_diff /method = enter comm_diff comm_sum.
```

**Coefficients**[a]

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | -.981 | .388 | | -2.527 | .015 |
| | comm_diff | .590 | .135 | .526 | 4.385 | .000 |
| | comm_sum | -.275 | .216 | -.153 | -1.272 | .210 |

a. Dependent Variable: int_diff

Too Hand Wavey? Try this! Estimate the model of Y in each condition then average the two regression parameters.

```
regression dep = int_G /method = enter comm_G.
```

```
regression dep = int_I /method = enter comm_I.
```

## Analysis using Judd et al. (2001)

4. Does the difference in communal goal affordance account for all the difference in interest?

Intercept of the regression for Step 3 provides an answer to this question.

What is the predicted difference in *Y*'s when there are no differences in *M*'s.

With the data: Is there a significance difference in interest predicted when there is no difference in communal goals?

**Coefficients**[a]

| Model | | Unstandardized Coefficients | | Standardized Coefficients | | |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | t | Sig. |
| 1 | (Constant) | -.981 | .388 | | -2.527 | .015 |
| | comm_diff | .590 | .135 | .526 | 4.385 | .000 |
| | comm_sum | -.275 | .216 | -.153 | -1.272 | .210 |

a. Dependent Variable: int_diff

31

## Analysis using Judd et al. (2001)

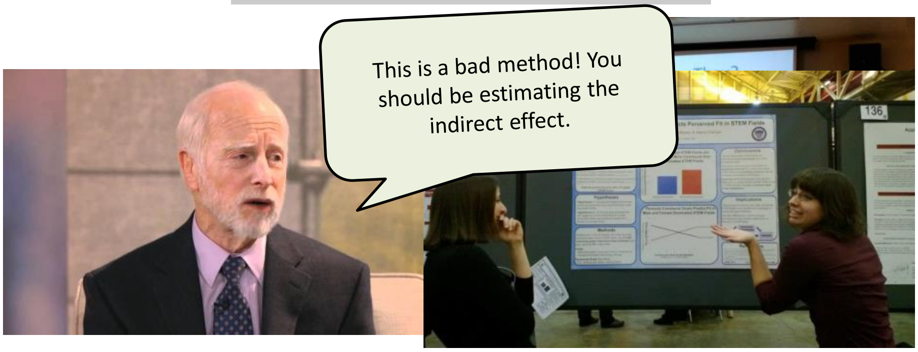
1. On average, is interest higher in the group work condition?
2. On average, is communal goal affordance higher in the group work condition?
3. Does difference in communal affordance predict a difference in interest?
4. Does the difference in communal goal affordance account for all the difference in interest?

**According to Judd, Kenny, and McClelland we do not have a mediated effect!**

Because there is no evidence that interest is higher in the group work condition, the Judd et al. (2001) method would conclude there is not mediation.

32

## No Mediation ☹




**Within-Subjects Version (CASC_WS.sps) :**
Female participants (N = 51) read two syllabi for a different computer science classes. One of the syllabi reported the class would have group projects throughout, and the other syllabi stated that individual project would be scheduled throughout.

**Measured Variables:**
- Interest in each the class (aggregate over 4 items)
- Perceptions that the class has a communal environment (aggregate over 5 items)

## Judd et al. Criticisms and Misuses

All criticisms of the causal steps approach apply to this approach:

- There is no explicit quantification of the indirect effect
  - Inference about an indirect effect should be the result of a test on a *quantification* of the indirect effect
- Requiring that there must be a total effect is too restrictive
  - The direct and indirect effect could be of opposite sign
  - There is greater power to detect the indirect effect than direct effect *(Judd, Kenny, 2014, Psych Science)*

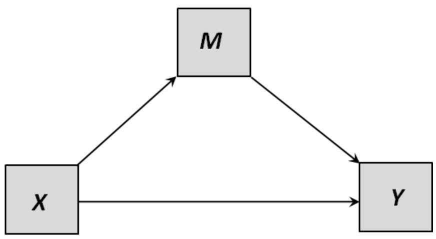
This method has been used by a variety of researchers:

- Approximately 300 citing papers, with around 140 using this method
- Many researchers do not report or estimate the partial regression coefficient for the sum of the mediators
- Because the estimate of the indirect effect is not made explicit, researchers often misinterpret the coefficients
  - $b_1$ path is often interpreted as indirect effect
- Extensions to more complicated models have been poorly implemented

34

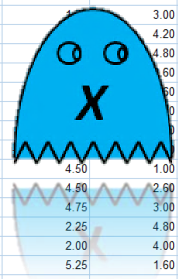
## Can we think about it like a path analysis?

**Analytic Goal: Can <u>group work</u> in computer science classes increase women's <u>interest</u> by increasing their perception that computer science is <u>communal</u>?**




Where is *X* in the data?

| $Y_1$ | $Y_2$ | $M_1$ | $M_2$ |
|---|---|---|---|
| int_I | int_G | comm_I | comm_G |
| 1.50 | 4.00 | 1.00 | 6.80 |
| 2.75 | 3.25 | 2.00 | 5.40 |
| 5.75 | 2.50 | 3.20 | 3.60 |
| 3.50 | 5.75 | 1.60 | 5.20 |
| 2.25 | 2.00 | 4.40 | 4.60 |
| 1.50 | [illegible] | 3.00 | 5.00 |
| 2.50 | [illegible] | 4.20 | 4.40 |
| 6.00 | [illegible] | 4.80 | 2.40 |
| 3.00 | [illegible] | [illegible] | 5.80 |
| 4.00 | [illegible] | [illegible] | 5.00 |
| 5.00 | [illegible] | [illegible] | 6.20 |
| 2.00 | [illegible] | [illegible] | 4.20 |
| 1.00 | [illegible] | [illegible] | 3.20 |
| 1.25 | 4.50 | 1.00 | 6.00 |
| 5.75 | 4.50 | 2.60 | 6.00 |
| 3.25 | 4.75 | 3.00 | 6.20 |
| 2.75 | 2.25 | 4.80 | 4.60 |
| 5.50 | 2.00 | 4.00 | 7.00 |
| 1.75 | 5.25 | 1.60 | 5.60 |
| 4.00 | 5.50 | 1.80 | 5.40 |
| 2.25 | 4.00 | 2.20 | 4.80 |
| 4.00 | 6.50 | 2.00 | 6.80 |
| 5.00 | 4.50 | 3.20 | 6.00 |
| 5.00 | 3.75 | 4.00 | 4.80 |
| 4.75 | 5.25 | 1.20 | 6.60 |






## Advantages of a path analytic approach

Provides an estimate of the indirect, total, and direct effects
- Allows us to conduct inferential tests directly on an estimate of the indirect effect

Connects researchers understanding of between-subjects mediation to within-subjects mediation
- Reduce misinterpretation of regression coefficients

Using a path analytic framework will help extend the simple mediation model to more complicated questions
- Multiple mediators
- Moderated mediation
- Integration of between and within-subjects designs

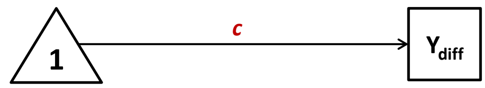
## Interpreting the Coefficients

**Total Effect (*c*):** The effect of our presumed cause (*X*) on our outcome (*Y*), without controlling for any other variables.

***a*-path:** The effect of our presumed cause (*X*) on our mediator *(M).*

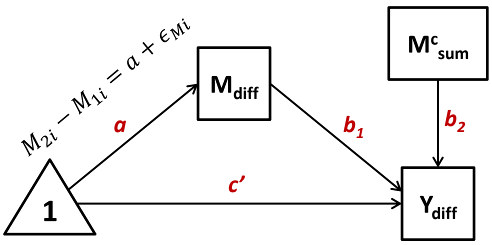
***b*-path:** The effect of our mediator (*M*) on the outcome (*Y*) while controlling for *X*. (i.e. predicted difference in *Y* for two people with the same score on X but who differ on *M* by one unit).

**Direct effect (*c'*):** The effect of our presumed cause (*X*) on *Y* while controlling for *M*. (i.e. predicted difference in *Y* for two people who differ by one unit on *X* but with the same score on *M*)




$$Y_{2i} - Y_{1i} = c + \epsilon_{Y*i}$$

[c] Indicates mean centered




$$Y_{2i} - Y_{1i} = c' + b_1(M_{2i} - M_{1i}) + b_2(M_{2i} + M_{1i} - \overline{(M_2 + M_1)}) + \epsilon_{Yi}$$

**Indirect Effect (*ab*):** Product of effect of *X* on *M*, and effect of *M* on *Y* controlling for *X*. The effect of *X* on *Y* through *M*.



## Within Subjects: Path Estimates

Total Effect c: (Regress $Y_D$ on a constant)

$$\widehat{Y_D} = c$$
$$\widehat{Y_D} = .373$$

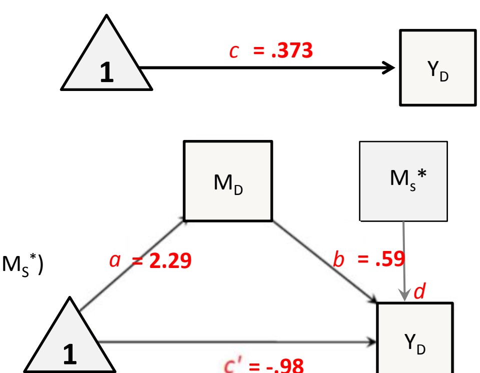
a path: (Regress $M_D$ on a constant)

$$\widehat{M_D} = a$$
$$\widehat{M_D} = 2.29$$

b path and c' path: (Regress $Y_D$ on $M_D$ and $M_S^*$)

$$\widehat{Y_D} = c' + bM_d + dM_S^* + e_3$$
$$\widehat{Y_D} = -.98 + .59M_D - .28M_S^*$$

A one unit increase in the difference in communal goal affordance is expected to result in a .59 unit increase in the difference in interest.
People with no difference in communal goal affordance perceptions are expected to be .98 units more interested in the individual class than the group work class .



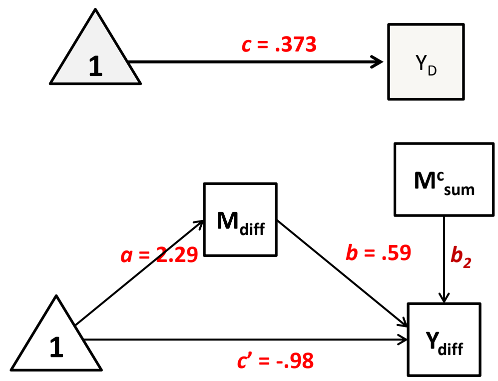

Note: $M_s$ must be mean centered for c' to have intended interpretation

## Data Example: Partitioning effect of X on Y

**The effect of *X on Y* partitions into two components: direct and indirect, in the usual way.**

| | | | | | | |
|---|---|---|---|---|---|---|
| c | = | c' | + | a | x | b |
| .373 | = | -.98 | + | 2.29 | x | .59 |
| .373 | = | -.98 | + | | 1.35 | |

**We can conduct inferential tests on the estimate of the indirect effect as in any other mediation analysis.**

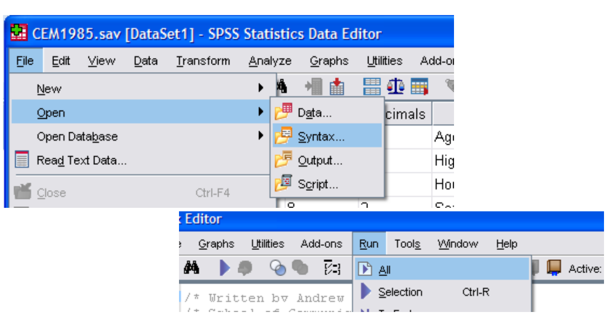



MEMORE has three methods of inference for the indirect effect available: bootstrapping, Monte Carlo confidence intervals, Sobel Tests



## Teaching your package MEMORE

MEMORE is a command which must be taught and re-taught to your statistical package (SPSS) every time you open the package. To teach your program the MEMORE command, open the memore.sps file and run the script exactly as is.




SPSS now knows a new command called MEMORE

# Writing MEMORE Syntax

PROCESS has 2 required arguments: `Y` and `M`

```
MEMORE m= comm_G comm_I /y = int_G int_I /normal=1/samples=10000
    /conf = 90.
```

**M** is your list of mediators (order matters)
**Y** is you list of outcomes (order should be matched to the order in the M list)
Some other arguments:

- **normal = 1** asks for Sobel test
- **samples** corresponds to the number of bootstrap/MC samples you would like
- **conf** specifies level of confidence you want (default is 95)
- **mc = 1** asks for Monte Carlo confidence intervals
- **bc = 1** asks for bias corrected bootstrap confidence intervals
- **conf** specifies level of confidence you want (default is 95)



# Using MEMORE for CASC WS data

```
MEMORE m= comm_G comm_I /y = int_G int_I.
```

```
******************* MEMORE Procedure for SPSS Version 1.1 ********************

                          Written by Amanda Montoya

                    Documentation available at afhayes.com

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

Variables:
Y = int_G    int_I
M = comm_G   comm_I

Computed Variables:
Ydiff =          int_G     -       int_I
Mdiff =          comm_G    -       comm_I
Mavg =  (        comm_G    +       comm_I   )       /2       Centered

Sample Size:
  51

******************************************************************************
Outcome: Ydiff =  int_G     -       int_I

Model
        Effect         SE          t         df          p       LLCI       ULCI
'X'      .3725      .2795     1.3330    50.0000      .1886     -.1888      .9339
******************************************************************************
Outcome: Mdiff =  comm_G    -       comm_I

Model
        Effect         SE          t         df          p       LLCI       ULCI
'X'     2.2941      .2491     9.2108    50.0000      .0000     1.7938     2.7944
******************************************************************************
```

Outcome variable

*c* = .37

*a* = 2.29

First part of output repeats what you told MEMORE to do. Always double check that this is correct!

Each chunk of the output like this is one of the regression models involved in the mediation analysis. This is the model of *M* from *X*, therefore this is the model which produces the estimate of *a*

## Using MEMORE for CASC WS data

```
MEMORE m= comm_G comm_I /y = int_G int_I.
```

```
**************************************************************************
Outcome: Ydiff =  int_G     -       int_I

Model Summary
          R       R-sq       MSE          F        df1        df2          p
      .5639      .3180    2.8299    11.1909     2.0000    48.0000      .0001

Model
            coeff         SE          t         df          p       LLCI       ULCI
'X'        -.9814      .3884    -2.5269    48.0000      .0149    -1.7623     -.2005
Mdiff       .5902      .1346     4.3845    48.0000      .0001      .3195      .8608
Mavg       -.5505      .4328    -1.2718    48.0000      .2096    -1.4208      .3198

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

Total effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
      .3725      .2795     1.3330    50.0000      .1886     -.1888      .9339

Direct effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
     -.9814      .3884    -2.5269    48.0000      .0149    -1.7623     -.2005

Indirect Effect of X on Y through M
             Effect     BootSE   BootLLCI   BootULCI
Ind1         1.3540      .3260      .6827     1.9653

Indirect Key
Ind1   X        ->      M1diff    ->      Ydiff
```

$c' = -.98$

$b = .590$

This is the model predicting $Y_{diff}$ from a constant, $M_{diff}$, and $M^{c}_{avg}$ therefore this model gives us an estimate of *b* and *c'*

Important effects for mediation and inference about these effects

Based on a bootstrap confidence interval we have evidence of mediation!



## Writing up a Mediation Analysis

**Is the effect of protesting on liking mediated by perceptions of response appropriateness?**

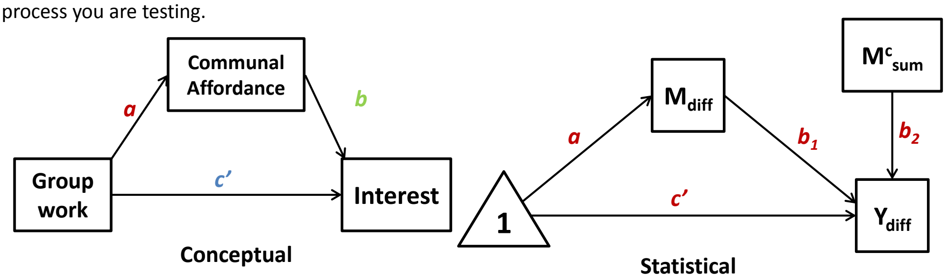
Overall there was no evidence of a total effect of group work on interest in computer science classes, we estimate that individuals were .37 units higher on interest in group work than individual work classes (*p* = .19). The class with group work was rated 2.29 units higher on communal goal affordance than the class with individual work (*p* < .001). A one unit increase in perception of communal goal affordance increased interest in the class by .59 units (*p* = .0001), and the relationship between communal goal affordance and interest in a class did not depend on condition (*p* = .21). The effect of group work on interest through communal goal fulfillment was different from zero (*ab* = 1.35, 95% Bootstrap CI [.68, 1.96]). This means that we expect women to be 1.35 units more interested in a computer science class with group work compared to one without group work, through the effect of group work on communal goal affordance, and the subsequent effect of communal goal affordance on interest. There was a significant direct effect between group work and interest (*c'* = -.98, p = .01). This indicates that there may be some other process, separate from communal goal affordance, which is actually deterring women from computer science classes with group work.

Tips:

- Walk the reader through the steps of the mediation in a way that is intuitive.
  - Include interpretations of the results: "The total effect was significant, *p* < .05"
  - Use equations and numbers *where helpful*.
  - Avoid using computational variable names (e.g. RESPAPPR)
- Avoid causal language if it is not supported by your research design.
- Pick one inferential method and report it
- Read the write ups of other's mediation analyses

## Visualizations

I suggest using both a conceptual and statistical visualization in order to help the reader understand the process you are testing.




Tips:

- Providing a conceptual diagram helps the readers understand the process you are interested in.
- Providing a statistical diagram helps readers understand how you estimated the model, and that you did it correctly.
- Provide path estimates on statistical diagram or in a table.
- Don't forget to report the path estimates and statistics for the $b_2$ path. It's important!

45

## Common Questions

- Can this method be used for more than two conditions?
  YES! Judd, Kenny, and McClelland (2001) describe a system for setting up contrasts among conditions, and testing the indirect effects of those contrasts.
  I recommend reading Hayes & Preacher (2014) on mediation analysis with a multicategorical IV if you want to try this out. I am happy to give instructions on how to trick MEMORE into doing this.
  There will be functionality (soonish) for MEMORE to do this.

  ALTERNATIVES: Some of the other repeated-measures mediation options are more appropriate if you have more than two conditions, so take a look at those when thinking about these options.
- Can I use multiple mediators?
  YES! MEMORE is already set up to do parallel mediation with up to 10 sets of mediators and serial mediation with up to two sets of mediators (See Montoya & Hayes, in press for instructions).
- Can we do conditional process models?
  Not yet, but we're working on it.
- How do I control for covariates?
  All of MEMORE's analysis are within-person models, so you do not need to control for any between subjects variables such as age, gender, big-5.
  Sometimes there are covariates which change within a person across conditions that you want to account for, this can be done by treating this additional variable as another set of mediators.

46

## Using MEMORE for CASC WS data

**Do people just like group work classes because they are easier?**

```
MEMORE m = comm_I comm_G diff_I diff_G /y = int_I int_G.
```

```
**************************************************************************
Outcome: Ydiff =  int_I     -       int_G

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .6307      .3978     2.6073     7.5978     4.0000    46.0000      .0001

Model
             coeff         SE          t         df          p       LLCI       ULCI
'X'          .9172      .3815     2.4042    46.0000      .0203      .1493     1.6851
M1diff       .4847      .1448     3.3460    46.0000      .0016      .1931      .7762
M2diff      -.4123      .1878    -2.1952    46.0000      .0332     -.7904     -.0342
M1avg        .5160      .4157     1.2411    46.0000      .2209     -.3209     1.3528
M2avg       -.3781      .2879    -1.3133    46.0000      .1956     -.9577      .2014

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

Total effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
     -.3725      .2795    -1.3330    50.0000      .1886     -.9339      .1888

Direct effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
      .9172      .3815     2.4042    46.0000      .0203      .1493     1.6851

Indirect Effect of X on Y through M
          Effect     BootSE   BootLLCI   BootULCI
Ind1     -1.1119      .3812    -1.8531     -.3522
Ind2      -.1779      .1160     -.4465      .0000
Total    -1.2897      .3507    -1.9566     -.5612

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

Notice that we are now **controlling** for difficulty of the class when estimating the effect of communal goal affordance on interest!

Controlling for difficulty, there is still a significant indirect effect through communal affordance

47

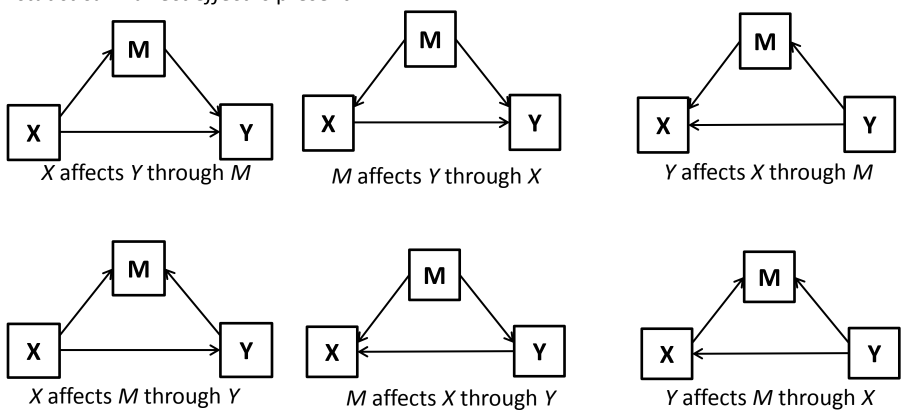
## Other Types of Repeated Measures Mediation

- Multilevel Models (Nesting is nuisance)
  - Bauer, Preacher, Gil (2006) *Psychological Methods*
    Covers Mediation and Moderated Mediation for 1-1-1 multilevel mediation
  - Kenny, Korchmaros, Bolger (2003) *Psychological Methods*
    Covers mediation for 1-1-1 multilevel models
- Latent Growth Curve Models (Longitudinal Processes M-Y measured over time)
  - Choeng, MacKinnon, Khoo (2003) *Structural Equation Modeling*
- Structural Equation Modeling (Can be used for a variety of data types)
  - Cole & Maxwell (2003) *Journal of Abnormal Psychology*
    X, M, and Y all measured over time
  - Newsom (2009) *Structural Equation Modeling*
    Dyadic data using LGMs
  - Selig & Little (2012) *Handbook of Developmental Research Methods*
    - Autoregressive models and cross-lagged panel models for longitudinal data X, M, and Y all measured over time.
- **Selig & Preacher (2009) *Research in Human Development***
  - **Longitudinal Models X, M, and Y measured across time. Cross-lagged panel models, latent growth models, latent difference score models**
- Multilevel SEM
  - Preacher, Zyphyr, Zhang, 2010
  - Preacher, Zhang, Zyphur, 2011

48

## A Brief Caution on Causality

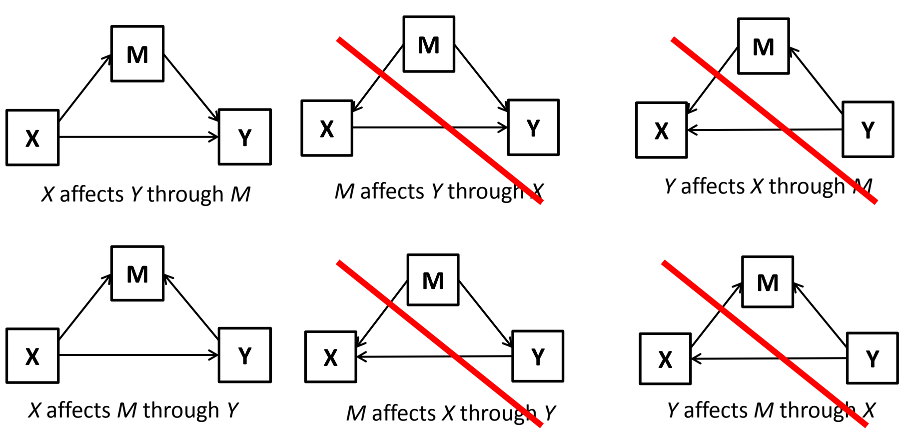
There are a number of alternative causal processes that may be occurring when a *statistical indirect effect* is present




49

## A Brief Caution on Causality

**What you get by manipulating *X*.**

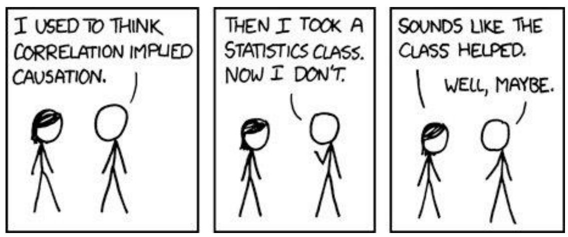



Even when *X* is manipulated, we can not provide evidence for the causal order between *M* and *Y*. This can only be supported using other experiments or previous research. *A statistically significant indirect effect does not lend credence to one model over another.*

50

## Thank you!

I am available for questions after the workshop and via email at montoya.29@osu.edu
I also live in LZ 216.




51

## Other Kinds of Bootstrap Confidence Intervals

All bootstrap confidence intervals use the same basic sampling technique, just use different methods for choosing the end points of the confidence intervals

**Bias-Corrected Confidence Interval**

- Percentile bootstrapping assumes that your sample estimate (*ab*) is unbiased in estimating the population indirect effect
- Bias-corrected reduces this assumption to assuming that the bias of *ab* is a constant (i.e. as N goes to infinity *ab* will go to the population indirect effect plus some constant)
- Bias-corrected confidence intervals estimate the bias of *ab* then adjust edges of confidence interval to be "bias-corrected" (i.e. centered not around your original estimate of *ab*), but around the point based on the bias estimation.

**Bias-Corrected and Accelerated**

- Same principles as BC regarding bias correction
- Acceleration allows for the assumption that the standard error of the indirect effect depends on the population value of the indirect effect
- Acceleration parameter, which is used to adjust the ends of the confidence interval is estimated using leave-one-out estimates of skew of the estimates of the indirect effect.

52