Mediation, Moderation, and Conditional Process Analysis I

Instructor: Amanda K. Montoya

GLOBAL
SCHOOL
EMPIRICAL
RESEARCH
METHODS

8-12 January 2018
Global School in Empirical Research Methods, Oslo, Norway

Understanding causal effects

Hagtvedt, H., & Patrick, V. M. (2008). Art infusion: The influence of visual art on the perception and evaluation of consumer products. *Journal of Marketing Research, 45*, 379-389.

HENRIK HAGTVEDT and VANESSA M. PATRICK*

In this research, the authors investigate the phenomenon of "art infusion," in which the presence of visual art has a favorable influence on the evaluation of consumer products through a content-independent spillover of luxury perceptions. In three studies, the authors demonstrate the art infusion phenomenon in both real-world and controlled environments using a variety of stimuli in the contexts of packaging, advertising, and product design.

Keywords: visual art, luxury, aesthetics, spillover effects, packaging, advertising, product design

Art Infusion: The Influence of Visual Art on the Perception and Evaluation of Consumer Products

How does the presence of visual art alter the way people view a consumer product? Throughout history, art has had the ability to arouse the imagination and capture the attention. Therefore, it is not surprising that art images are often used to promote unrelated products—for example, by being displayed in advertisements (Hetsroni and Tukachinsky 2005). It is proposed that such "high-culture images reach more people more often through advertising than through any other medium" (Hoffman 2002, p. 6). Other times, art becomes an integrated part of a product, such as when furniture is artistically designed or a painting is printed on a shirt. Some companies, such as De Beers, use art in image promotion, conveying the idea that diamonds, like paintings, are unique works of art (Epstein 1982). Sometimes, art is even created for the sole purpose of marketing a product, such as in the enduring Absolut Vodka advertising campaign (Lewis 1996).

It is clear that influential marketing practitioners believe that art somehow has the power to influence consumer perceptions. Vast amounts of money are spent on representing visual art in conjunction with products, in the hope that the products will become more marketable as a result. However, the issue of whether these beliefs are well founded remains unresolved. Furthermore, there is little evidence to suggest that marketing professionals have been provided with the scientific basis necessary to use visual art in a strategic manner rather than purely on the basis of experience and intuition. Supplying this basis is a complex endeavor. However, the current research represents an initial step to analyze systematically the influence of visual art on consumer evaluations of the products with which it is associated. This influence represents a fundamental gap in current understanding, not only in terms of the $23.5 billion global art market (Kusin & Company 2002) but also in terms of the potential impact of art on other markets and marketing activities.

In this research, we examine the phenomenon of "art infusion," which we broadly define as the general influence of the presence of art on consumer perceptions and evaluations of products with which it is associated. More specifically, we theorize that perceptions of luxury associated with visual art spill over from the artwork onto products with which it is associated, leading to more favorable evaluations of these products. Furthermore, we propose that this influence does not depend on the content of the specific artwork—that is, what is depicted in the artwork—but rather on general connotations of luxury associated with visual art.

In Study 1, we demonstrate the art infusion phenomenon in a real-world setting. In this study, consumers are briefly exposed to art or nonart images, which are matched for

*Henrik Hagtvedt is a doctoral student in Marketing (e-mail: hhagtvedt@yahoo.com), and Vanessa M. Patrick is Assistant Professor of Marketing (e-mail: vpatrick@uga.edu), Terry College of Business, University of Georgia. Both authors contributed equally to this research. They express their appreciation to the anonymous *JMR* reviewers for their thoughtful comments and encouragement and to the marketing faculty at the University of Georgia for their support of this research. They also thank Reidar Hagtvedt, John Mowen, Gavan Fitzsimons, Shashi Matta, Piyush Kumar, and Srinivas K. Reddy for their input and Nicole Patrick, Simone Patrick, and Colin Heacock for their help in creating some of the stimuli used in this article. Chris Janiszewski served as associate editor for this article.

ISSN: 0022-2437 (print), 1547-7193 (electronic)
379
Journal of Marketing Research
Vol. XLV (June 2008), 379–389

Understanding causal effects

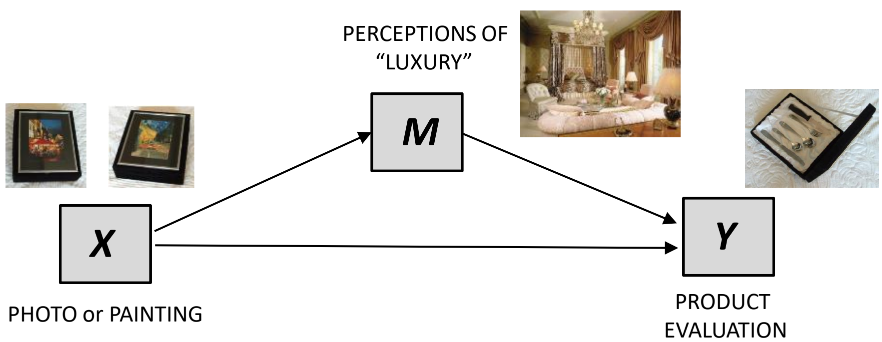
Hagtvedt, H., & Patrick, V. M. (2008). Art infusion: The influence of visual art on the perception and evaluation of consumer products. *Journal of Marketing Research, 45*, 379-389.

The product with the impressionist painting on the box was evaluated more favorably than the product with a photograph of a similar scene.

Remaining Questions:
- How does this effect occur? What is the *mechanism* that produces it?
- Is the effect consistent across type of product, type of consumer, and so forth.

These are the important questions!

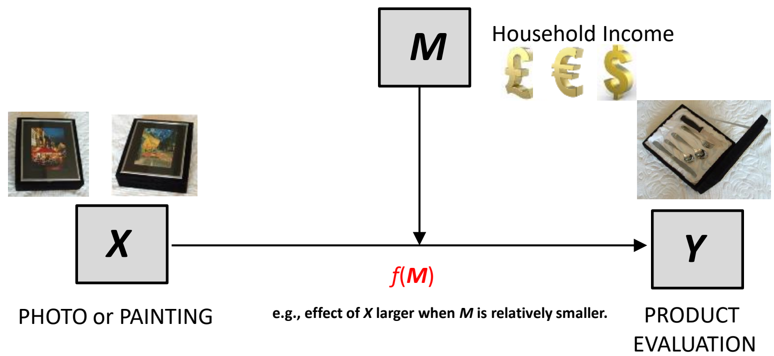
A simple mediation model



Art-infused product was perceived as more "luxurious," and this greater perceived luxury translated into a more favorable product evaluation. Thus, the infusion of the product with art influenced product evaluation at least partly through the "mechanism" of perceived luxuriousness.

Mediation analysis is about estimating and making inferences about such *indirect effects*.

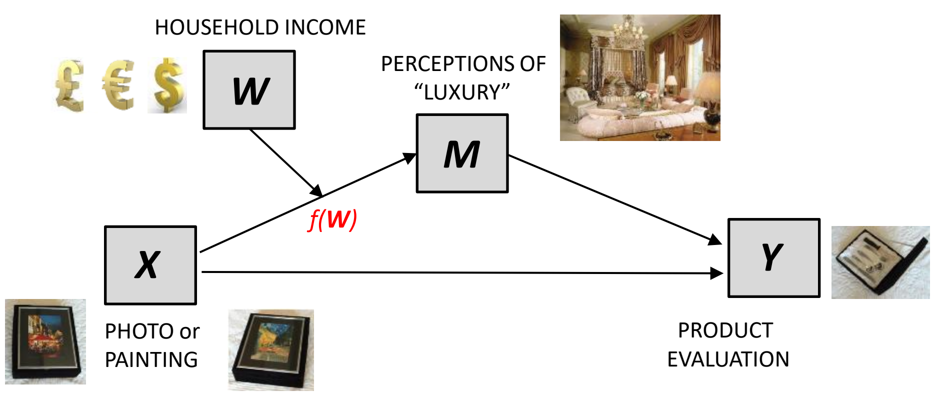
Moderation



Is this effect of art infusion is larger among those with less income? Is the size of the effect dependent on (or a function of) income. In this case, income is a *moderator* of the effect of art infusion on product evaluation.

Moderation analysis is about the estimation of *contingent effects*, i.e., examining the boundary conditions of effects or the factors that make effects large versus small, positive vs. negative.

Combining moderation and mediation



Does art infusion result in greater perceptions of luxury more so among those with less income? If so, then the indirect effect of art infusion on product evaluation through perceptions of luxury depends on income. Thus, the strength of this "mechanism" may depend on income. <u>Mediation</u> can be <u>moderated</u>.

In this class

- After a review of OLS regression, we start with questions of "**HOW**"—
  statistical mediation analysis

  "Direct," "indirect," and "total effects" in path models and how to test hypotheses in such models using OLS regression and various computational tools developed for this purpose.

- We then move to questions of "**WHEN**"--**moderation analysis**

  Estimation and interpretation of models in which a predictor can have different effects on an outcome depending on the value of another variable in the model.

- We then explore models that combine moderation and mediation—
  "conditional process analysis"

- With fundamentals covered, we address more complex models and issues in mediation and moderation analysis, such as multiple mediators and multi-categorical independent variables and moderators.

What you'll need

- **This course is hands-on. Hopefully you brought a laptop with SPSS 19+ or SAS 9.1+ with PROC IML. If not, that is ok. You'll still benefit.**

  SPSS Code   SAS Code

- **Various files available on a USB drive I pass around.**
  - SPSS and SAS data folders. SPSS data files are ready to go. SAS files are programs thus must be executed to make them "work" files.
  - SPSS and SAS PROCESS folders. This contains the PROCESS macro we'll heavily rely on, and some documents related to it.
  - Miscellaneous folder. Various files, including some PDFs and other miscellaneous things of relevance to this course.

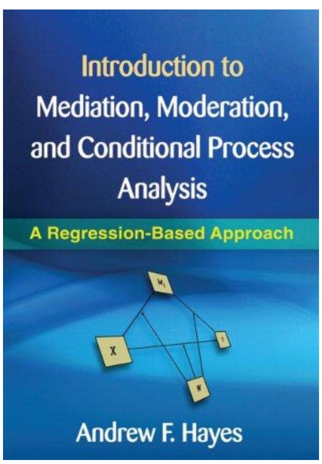
- **A lot of stamina.**

What we will and won't do

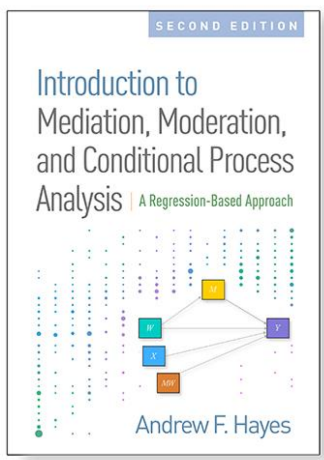
- We will stick with fairly simple models to cover basic principles, with continuous outcomes, and cross-sectional or experimental data.
- Statistical mediation analysis. No discussion of counterfactuals, "potential outcomes," directed graphs, or other approaches to thinking about cause.
- Everything OLS-regression-based.
- No dichotomous outcomes, nothing multilevel.

Although the principles are not software specific, their implementation is facilitated with the use of a "macro" which makes otherwise tedious things very simple and effortless. You will learn about PROCESS.

This course is a companion to…



2013



December 2017

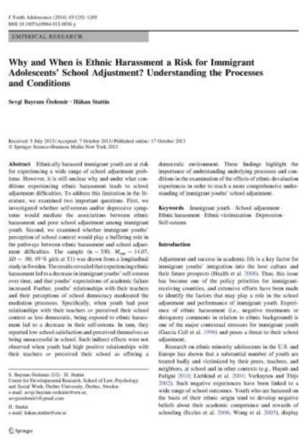
www.guilford.com

Example for the class “inspired by”…

Bayram-Ozdemir, S. & Stattin, H. (2014). Why and when is ethnic harassment a risk for immigrant adolescents’ school adjustment? Understanding the processes and conditions. ***Journal of Youth and Adolescence, 43***, 1252-1265.

Example for the class “inspired by…”

Bayram-Ozdemir, S. & Stattin, H. (2014). Why and when is ethnic harassment a risk for immigrant adolescents’ school adjustment? Understanding the processes and conditions. ***Journal of Youth and Adolescence, 43***, 1252-1265.



330 7th to 9th grade students in Finland measured in the spring term (T1) and again one year later (T2).

All were first or second generation immigrants who reported at least some ethnic harassment.

HARASS: 6-item measure of ethnicity-related harassment frequency (scaled 1 to 5). T1 only.

POSREL: 6-item measure of positivity of relationships with teachers (scaled 1 to 4). T1 only.

SE: 10-item Rosenberg self-esteem scale (scaled 1 to 4). T1 and T2.

DEP: 20-item Center for Epidemiological Studies Depression Scale for Children (scaled 1 to 4). T1 and T2.

FAIL: 4-item measure of perceived academic failure at school (scaled 1 to 4). T1 and T2.

SATIS: 5-item measure of satisfaction in school (scaled 1 to 5). T1 and T2.

All are continuous variables scaled such that higher = more.

The Data: HARASS

SPSS

*harass.sav [DataSet3] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

	harass	se2	dep2	satis2	fail2	posrel	se1	dep1	satis1	fail1
1	2.16	3.80	1.25	4.80	1.00	1.66	3.60	1.10	4.40	2.00
2	2.33	2.90	1.95	4.20	1.75	2.83	2.80	2.25	3.60	2.25
3	1.16	3.80	1.95	5.00	1.50	3.16	3.80	1.95	4.80	1.00
4	2.50	2.50	2.15	3.00	2.00	2.00	3.10	1.25	3.20	2.00
5	1.33	3.00	1.90	2.60	1.00	2.83				
6	1.16	2.50	1.25	3.40	1.25	2.83				
7	1.50	2.10	2.05	2.40	1.50	3.00				
8	1.50	3.30	1.75	2.80	1.25	3.00				
9	1.83	2.70	2.15	3.20	1.50	3.50				
10	1.33	2.80	2.45	4.00	2.00	3.16				

"1" and "2" in the variable name refers to time 1 and time 2, respectively. The absence of a number means the data were available only at time 1.

SAS

harass

```
data harass;
  input harass se2 dep2 satis2 fail2 posrel se1 dep1 satis1 fail1;
datalines;
2.16 3.80 1.25 4.80 1.00 1.66 3.60 1.10 4.40 2.00
2.33 2.90 1.95 4.20 1.75 2.83 2.80 2.25 3.60 2.25
1.16 3.80 1.95 5.00 1.50 3.16 3.80 1.95 4.80 1.00
2.50 2.50 2.15 3.00 2.00 2.00 3.10 1.25 3.20 2.00
1.33 3.00 1.90 2.60 1.00 2.83 2.60 1.55 2.40 1.75
1.16 2.50 1.25 3.40 1.25 2.83 2.80 1.45 3.00 2.00
1.50 2.10 2.05 2.40 1.50 3.00 1.80 2.10 2.40 2.00
1.50 3.30 1.75 2.80 1.25 3.00 3.40 1.55 3.80 1.25
1.83 2.70 2.15 3.20 1.50 3.50 2.90 2.20 3.40 1.50
1.33 2.80 2.45 4.00 2.00 3.16 2.80 1.25 4.00 1.25
1.33 2.60 1.70 3.60 1.75 3.16 3.00 2.05 4.40 2.25
```

The SPSS file is ready for analysis. The SAS version is a SAS program that must be executed to produce a temporary work data file.

These are not the actual data from this study. They were generated to produce similar results to the published study.

A quick review of regression analysis

- Linear regression is the foundation of this class.
- Used throughout science as a means of "modeling" the relationship between variables.
- Many of the kinds of analyses and statistics you already know about can be expressed in the form of a linear regression model
  - independent groups *t* test
  - analysis of variance

Pearson's coefficient of correlation (*r*) is the building block of linear regression analysis. Consider the correlation between positivity of relationships with teachers and satisfaction at school (measured contemporaneously).

Correlations

		posrel	satis1
posrel	Pearson Correlation	1	.458
	Sig. (2-tailed)		.000
	N	330	330
satis1	Pearson Correlation	.458	1
	Sig. (2-tailed)	.000	
	N	330	330

SPSS code in black box

```
correlations variables = posrel satis1.
```

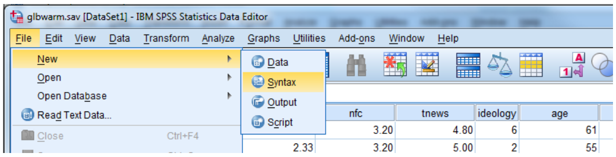
SAS code in white box

```
proc corr data=harass;var posrel satis1;run;
```

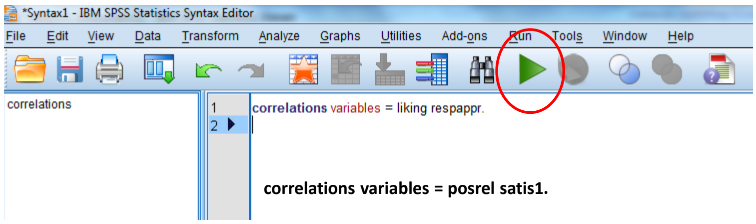
Using SPSS syntax

We will use syntax to instruct SPSS what to do in this class. There are many benefits of learning how to write SPSS syntax.

(1) Open a new syntax window (File > New > Syntax)



(2) Type your command(s) into the blank window that opens



(3) Click and drag to highlight code you want to execute and press the "play" button or select various options under "Run" in the syntax window menu.

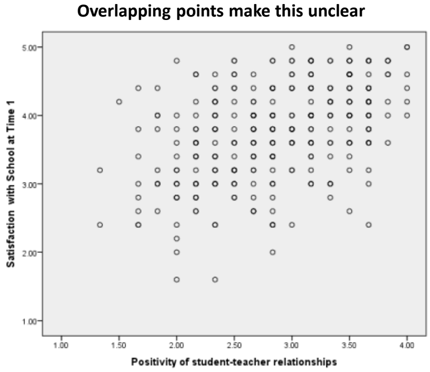
A scatterplot

Consider a scatterplot visually depicting this relationship:

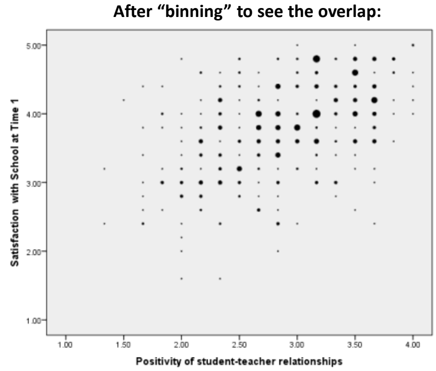
```
graph/scatterplot=posrel with satis1.
```

```
proc sgscatter data=harass;plot satis1*posrel;run;
```

Overlapping points make this unclear



After "binning" to see the overlap:



If you had to draw a single straight line through this plot that "best fits" the relationship, where would you draw it? At its heart, this is the problem regression analysis solves.

OLS (Ordinary Least Squares) linear regression

Goal: Derive the equation ("model") for the line representing the association between independent variable *X* and dependent variable *Y* that "best fits" the data.

The "simple regression model" (i.e., only one variable on the right hand side) takes the form

$$Y_j = i + bX_j + e_j$$

Using the ordinary least squares criterion, there is only one line described by the function

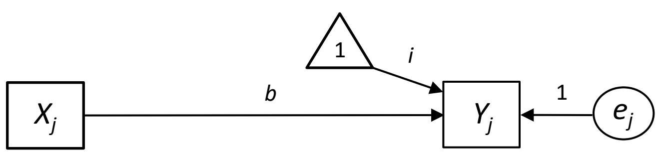
$$\hat{Y}_j = i + bX_j \qquad e_j = Y_j - \hat{Y}_j$$

that "best fits" the data, where $\hat{Y}_j$ is the estimated or fitted value of *Y*, and "best fit" is defined as the line that minimizes the sum of the squared residuals ($SS_{residual}$), summed over all *n* cases in the data. This is called the LEAST SQUARES criterion.

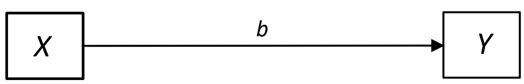
$$SS_{residual} = \sum_{j=1}^{n} (Y_j - \hat{Y}_j)^2 = \sum_{j=1}^{n} e_j^2$$

A visual representation

$$Y_j = i + bX_j + e_j$$



or, in shorthand,



In a diagram such as this, ⟶ represents "predictor of" or "component of" but not necessarily "cause of," although the association could be causal. So *X* is a predictor of *Y* (and *perhaps* a cause of *Y*) in this diagram.

Easier to do in SPSS and then explain

This is an easy problem for a computer with an OLS regression routine. We estimate Y from X, or regress Y on X. X = POSREL (positivity of teacher-student relationships), Y = SATIS1 (satisfaction with school).

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.458[a]	.210	.207	.61840

a. Predictors: (Constant), posrel

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	33.263	1	33.263	86.980	.000[b]
	Residual	$\boxed{125.433}$	328	.382		
	Total	158.696	329			

a. Dependent Variable: satis1

b. Predictors: (Constant), posrel

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B Lower Bound	95.0% Confidence Interval for B Upper Bound
1	(Constant)	2.245	.166		13.527	.000	1.918	2.571
X	posrel	.531	.057	.458	9.326	.000	.419	.643

a. Dependent Variable: satis1

$SS_{residual}$ = 125.433

No other *i*,*b* pair would produce a smaller value of $SS_{residual}$.

$$\hat{Y}_j = i + bX_j$$

$$\hat{Y}_j = 2.245 + 0.531X_j$$

This is the best fitting OLS regression model, assuming a linear association between X and Y.

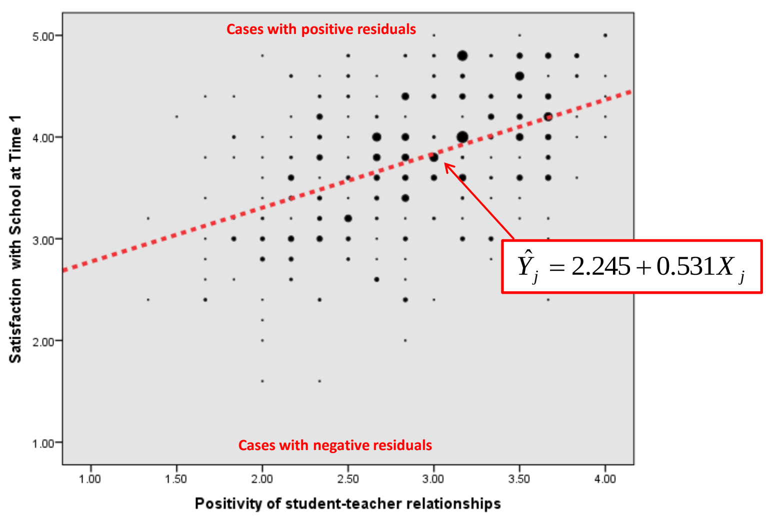
```
regression/statistics defaults ci/dep=satis1/method=enter posrel.
```

```
proc reg data=harass;model satis1 = posrel/stb clb;run;
```

Output A

The model in visual form

Y_j



X_j

Interpretation of *i* and *b*

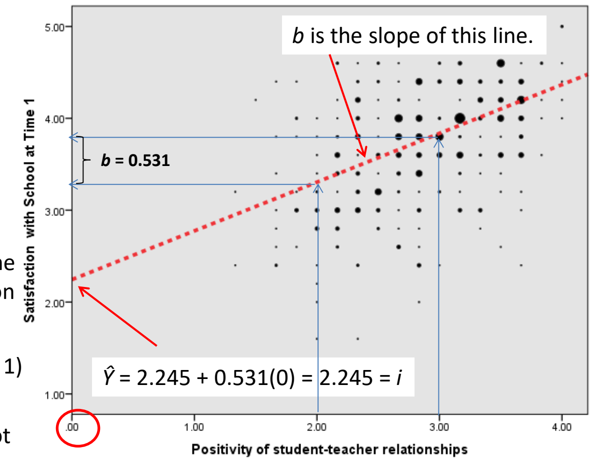
$$\hat{Y}_j = i + bX_j$$

$$\hat{Y}_j = 2.245 + 0.531X_j$$

b = estimated difference in *Y* between two cases that differ by one unit on *X*. The sign of *b* speaks to the sign of the association between *X* and *Y*.

$b = \hat{Y} | (X = \theta) - \hat{Y} | (X = \theta - 1)$

i = estimated value of *Y* when *X* = 0. This is not meaningful here.



Two kids that differ by one unit in the positivity of their student-teacher relationships are estimated to differ by *b* = 0.531 units in satisfaction with school. The kid that is **higher** in positivity of those relationships is estimated to be *more* satisfied (because *b* is positive)

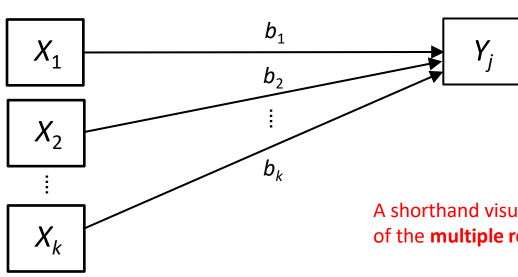
Multiple predictors

Multiple predictors variables are handled with ease, without modification to the estimation process. But this results in some interpretational changes.

$$Y_j = i + b_1X_{1j} + b_2X_{2j} + b_3X_{3j} + \ldots + b_kX_{kj} + e_j$$

or, more concisely,

$$Y_j = i + \sum_{m=1}^{k} b_m X_{mj} + e_j$$



A shorthand visual representation of the **multiple regression** model.

A multiple regression model (SPSS)

```
regression/statistics defaults ci/dep=satis1/method=enter posrel harass se1.
```

```
proc reg data=harass;model satis1 = posrel harass se1/stb clb;run;
```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.545[a]	.297	.290	.58503

a. Predictors: (Constant), se1, posrel, harass

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	47.120	3	15.707	45.891	.000[b]
	Residual	111.576	326	.342		
	Total	158.696	329			

a. Dependent Variable: satis1

b. Predictors: (Constant), se1, posrel, harass

$SS_{residual}$ = 111.576

No other *i* and set of *b* values would produce a smaller value of $SS_{residual}$.

X_1 = positivity of teacher-student relationships

X_2 = harassment frequency

X_3 = self-esteem.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	1.479	.309		4.788	.000	.871	2.087
	posrel	.472	.055	.408	8.519	.000	.363	.581
	harass	-.145	.090	-.078	-1.613	.108	-.322	.032
	se1	.373	.064	.276	5.828	.000	.247	.499

a. Dependent Variable: satis1

$$\hat{Y}_j = 1.479 + 0.472X_{1j} - 0.145X_{2j} + 0.373X_{3j}$$

Output B

The meaning of the values of *b*

$$\underbrace{\hat{Y}_j}_{\text{Satisfaction}} = 1.479 + \underbrace{0.472X_{1j}}_{\text{Positive relationships}} \underbrace{- 0.145X_{2j}}_{\text{harassment}} \underbrace{+ 0.373X_{3j}}_{\text{Self-esteem}}$$

Two kids *the same on all predictors except the positivity of their teacher relationships* (X_1) but who <u>differ by one unit in such positivity</u> will differ by b_1 = 0.472 units in estimated satisfaction with school ($\hat{Y}$).

> b_1, the partial regression coefficient for X_1, quantifies how <u>differences in positivity of the student-teacher relationship relates to differences in satisfaction with school</u> when *all other predictor variables in the model are held constant*, or "statistically controlling for" those other variables.

Two people the <u>same on all predictors except ethnic harassment (X_2)</u> but who *differ by one unit in such harassment* will differ by b_2 = 0.145 units in estimated satisfaction with school ($\hat{Y}$).

> b_2, the partial regression coefficient for X_2, quantifies how *differences in frequency in the experience of ethnic harassment relate to differences in school satisfaction* when <u>all other predictor variables in the model are held constant</u>, or "statistically controlling for" those other variables. The negative sign for b_2 means those who experience *more* harassment are estimated to be *less* satisfied with school.

Statistical inference

$$Y_j = i + b_1 X_{1j} + b_2 X_{2j} + b_3 X_{3j} + \ldots + b_k X_{kj} + e_j$$

The constant, *i*, and *k* values of *b* are sample-specific. They are sample-specific estimates of a corresponding population or process model (the "true" model):

$$Y_j = {}_T i + {}_T b_1 X_{1j} + {}_T b_2 X_{2j} + \ldots + {}_T b_k X_{kj} + {}_T e_j$$

Departures between the "true model" and the obtained model resulting from our data are used to test hypotheses about the "true values" of *b*.

Departures between the true and the obtained model are assumed to be driven by "random" processes, such as random sampling, random assignment variation, measurement error, etc., unless the data suggest otherwise. We attempt to estimate the true model using our data, hoping that our estimates of the true values of that model are accurate.

Null hypothesis testing for ${}_T b$

$$Y_j = {}_T i + {}_T b_1 X_{1j} + {}_T b_2 X_{2j} + \ldots + {}_T b_k X_{kj} + {}_T e_j$$

In any study, we observe only b_j, the sample estimate of ${}_T b_i$. We often are interested in making an inference about the size of ${}_T b_i$, or testing a hypothesis about its value.

e.g., Null hypothesis test about ${}_T b_1$:

Assume ${}_T b_1$ equals some specific value. Typically, we assume ${}_T b_1 = 0$ under the **null hypothesis** (i.e., X_j is unrelated to *Y* when all other variables in the model are held constant).

$$H_0: \; {}_T b_1 = 0$$
$$H_a: \; {}_T b_1 \neq 0$$

If H_0 is true, then b_1 / s_{b1} follows the $t(df_{residual})$ distribution, where s_{b1} is the estimated standard error of b_1. Using the *t* distribution, we generate a *p*-value and reject H_0 in favor of H_a if $p \leq \alpha$-level chosen for the test (usually .05). In that case, the result is "statistically significant."

Statistical inference for partial regression coefficients

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	47.120	3	15.707	45.891	.000[b]
	Residual	111.576	326	.342		
	Total	158.696	329			

a. Dependent Variable: satis1

b. Predictors: (Constant), se1, posrel, harass

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	1.479	.309		4.788	.000	.871	2.087
	posrel	.472	.055	.408	8.519	.000	.363	.581
	harass	-.145	.090	-.078	-1.613	.108	-.322	.032
	se1	.373	.064	.276	5.828	.000	.247	.499

a. Dependent Variable: satis1

Output B

H_0: ${}_T b_1 = 0$

H_a: ${}_T b_1 \neq 0$

$b_1 = 0.472$, se(b_1) = 0.055, $t(326) = 8.519$, $p < 0.001$

Reject H_0 in favor of H_a

H_0: ${}_T b_2 = 0$

H_a: ${}_T b_2 \neq 0$

$b_2 = -0.145$, se(b_2) = 0.090, $t(326) = -1.613$, $p = 0.108$

Do not reject H_0

Two kids equal in ethnic harassment frequency and self esteem but who differ in the positivity of their student-teacher relationships differ from each other in satisfaction more than can be explained by chance. The observed and statistically significant positive partial relationship tells us that kids with more positive student-teacher relationships are more satisfied with school.

Two kids equal in the positivity of their student-teacher relationships and their self-esteem but who differ in ethnic harassment frequency do not differ in their satisfaction any more than would be expected by "chance."

Interval estimation

Inferences can also be framed as an interval such that this interval will capture the true value a certain percentage of the time. As a rough rule-of-thumb, we can be 95% confident that the true value resides within <u>about</u> 2 standard errors of the obtained estimate.

$$b_i - 2s_{b_i} \leq {}_T b_i \leq b_i + 2s_{b_i}$$

Output B

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	1.479	.309		4.788	.000	.871	2.087
	posrel	.472	.055	.408	8.519	.000	.363	.581
	harass	-.145	.090	-.078	-1.613	.108	-.322	.032
	se1	.373	.064	.276	5.828	.000	.247	.499

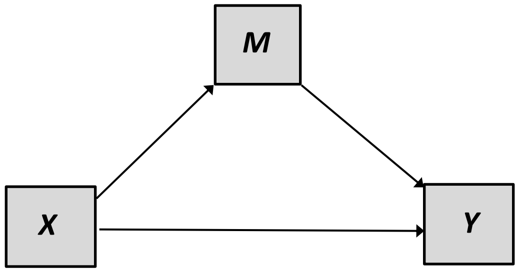
a. Dependent Variable: satis1

We can be 95% sure ${}_T b_1$ is somewhere between 0.363 and 0.581

We can be 95% sure ${}_T b_2$ is somewhere between -0.322 and 0.032.

Question: Is it fair to say we have 'no evidence' that kids who are harassed more frequently are less satisfied with school?

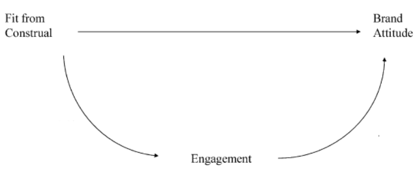
Statistical mediation analysis



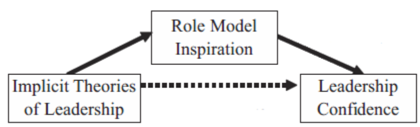
The "simple mediation" model

A mediation model links an assumed cause (*X*) to an assumed effect (*Y*) at least in part via an intermediary variable (*M*). An intermediary variable can be a psychological state, a cognitive process, an affective response, or any other conceivable "mechanism" through which *X* exerts an effect on *Y*. *X* affects *M* which in turn affects *Y*.

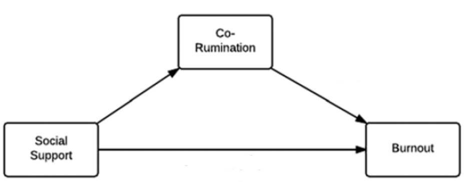
Some examples in the literature



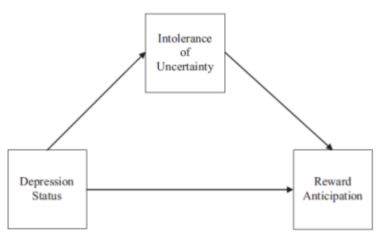
Lee, A. Y., Keller, P. A., & Sternthal, B. (2010). Value from regulatory construal fit: Persuasive impact of fit between consumer goals and message concreteness. ***Journal of Consumer Research***, *36*, 735-747.



Hoyt, C. L., Burnette, J. L., & Innella, A. N. (2012). I can do that: The impact of implicit theories on leadership model effectiveness. ***Personality and Social Psychology Bulletin***, *38*, 257-268.



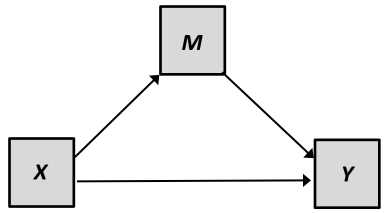
Boren, J. P. (2014). The relationship between co-rumination, social support, stress, and burnout among working adults. ***Management Communication Quarterly***, *28*, 3-25.



Nelson, B. D., Shankman, S. A., & Proudfit, G. H. (2014). Intolerance of uncertainty mediates reduced reward anticipation in major depressive disorder. ***Journal of Affective Disorders***, *158*, 108-113.

The most basic intervening variable model

- For *M* to be an intermediary, it must be located *causally between X* and *Y*.
- *M* is sometimes called a "mediator", but it goes by other names as well.

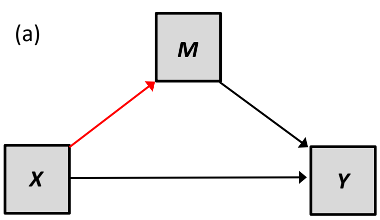


- Mediator models are causal models and carry with them the usual criteria for making causal claims.
    Difficult to establish cause statistically or otherwise.
    Theory is sometimes the sole foundation upon which our causal claims rest. That's ok so long as we recognize this.

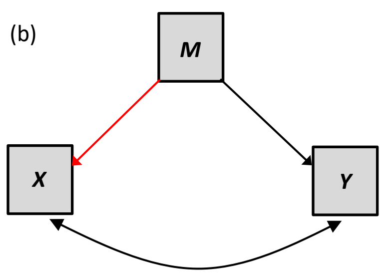
Mediation and spuriousness

Mediation analysis cannot distinguish between mediation (a) and spuriousness (b). If (b) can be deemed plausible, that weakens the case for (a) regardless of what the data analysis tells you.

(a)



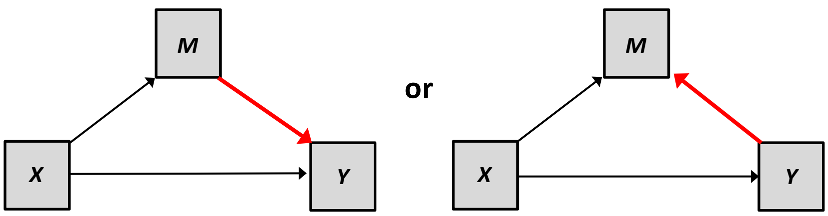
(b)



Inferences are always design-bound.
Mediation is a causal process, but causal claims are only justified if the data allow such claims, regardless of what the statistics say.

Causal order

Manipulation of and random assignment to *X* affords causal inference for the effect of *X* on *M* and *Y*, but not the effect of *M* on *Y*. We cannot establish causal order for the *M*-*Y* path using the methods that are the focus here. Theory is important. Multiple studies can help, one of which involves manipulation of *M*.



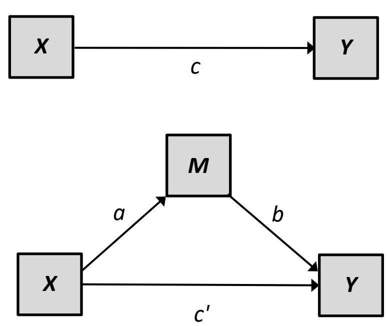
When *X* is not experimentally manipulated, all paths are subject to potential alternative causal orders.

Path analysis: Total, direct, and indirect effects

Let *a*, *b*, *c*, and *c'* be quantifications of causal effects, such as regression coefficients in an OLS model (or maximum likelihood path estimates in a structural equation model)

$$\hat{Y} = i_1 + cX$$
$$\hat{M} = i_2 + aX$$
$$\hat{Y} = i_3 + c'X + bM$$

A "simple mediation" model



c = "total effect" of *X* on *Y*

$a \times b$ = "indirect effect" of *X* on *Y*

c' = "direct effect" of *X* on *Y*

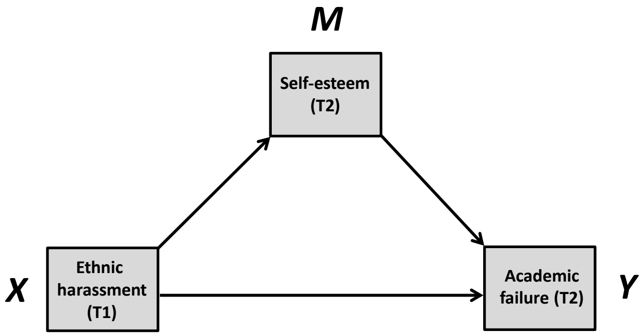
total effect = direct effect + indirect effect

$c \quad = \quad c' \quad + \quad (a \times b)$

indirect effect = total effect – direct effect

$(a \times b) \quad = \quad c \quad - \quad c'$

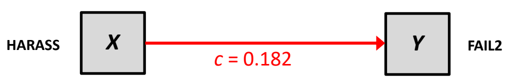
Our question



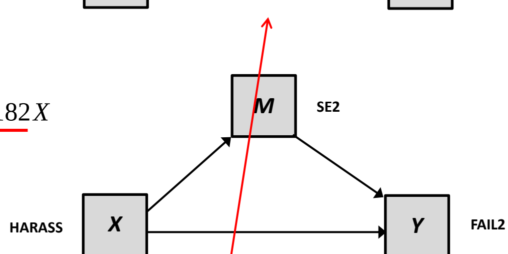
Does ethnic harassment influence school performance by affecting self-esteem which in turn affects performance.

This question is not asked contingent on evidence of simple association between *X* (ethnic harassment) and *Y* (performance)

Using a set of OLS regression analyses



$$\hat{Y} = 1.346 + \underline{0.182} X$$



Coefficients[a]

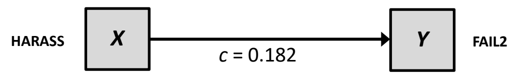
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.346	.113		11.866	.000
	harass	$\boxed{.182}$	.073	.136	2.478	.014

a. Dependent Variable: fail2

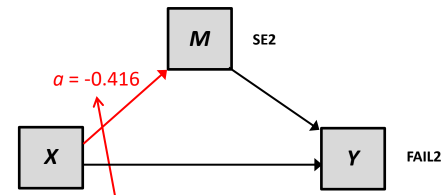
SPSS:
```
regression/dep=fail2/method=enter harass.
```

SAS:
```
proc reg data=harass;model fail2=harass;run;
```

Using a set of OLS regression analyses



$$\hat{M} = 3.597 - \underline{0.416}X$$



Coefficients[a]

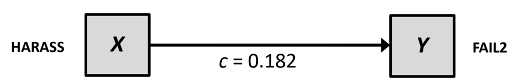
Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.597	.123		29.123	.000
	harass	$\boxed{-.416}$	.080	-.276	-5.209	.000

a. Dependent Variable: se2

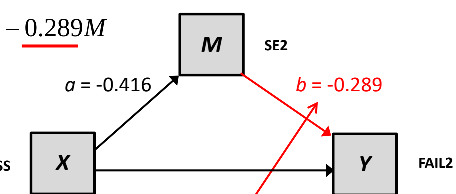
```
regression/dep=se2/method=enter harass.
```

```
proc reg data=harass;model se2=harass;run;
```

Using a set of OLS regression analyses



$$\hat{Y} = 2.385 + 0.062X - \underline{0.289}M$$



Coefficients[a]

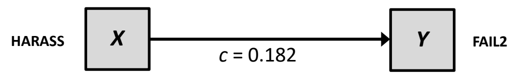
Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.385	.204		11.676	.000
	harass	.062	.072	.046	.850	.396
	se2	$\boxed{-.289}$	.048	-.324	-5.988	.000

a. Dependent Variable: fail2

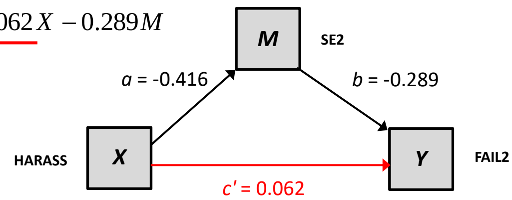
```
regression/dep=fail2/method=enter harass se2.
```

```
proc reg data=harass;model fail2=harass se2;run;
```

Using a set of OLS regression analyses



$$\hat{Y} = 2.385 + \underline{0.062}\,X - 0.289M$$



Coefficients[a]

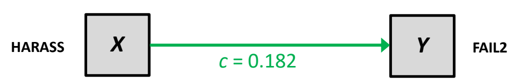
Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.385	.204		11.676	.000
	harass	$\boxed{.062}$	.072	.046	.850	.396
	se2	-.289	.048	-.324	-5.988	.000

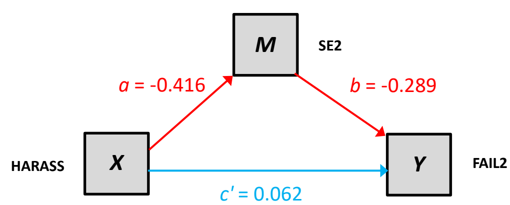
a. Dependent Variable: fail2

```
regression/dep=fail2/method=enter harass se2.
```

```
proc reg data=harass;model fail2=harass se2;run;
```

Using a set of OLS regression analyses





Direct effect of X on Y = c' = 0.062
Indirect effect of X on Y via M = ab = -0.416(-0.289) = 0.120
Total effect of X on Y = c' + ab = 0.062 + 0.120 = 0.182 = c

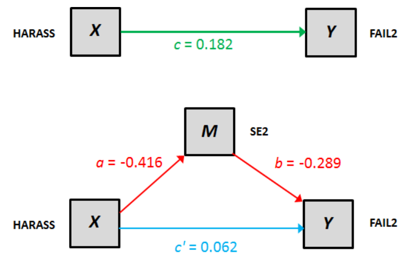
Interpretation of the total, direct, and indirect effects

Generic

Total: Two people who differ by one unit on *X* are estimated to differ by *c* units on *Y* on average.

Indirect: They differ by *ab* units on average as a result of the effect of *X* on *M* which in turn affects *Y*.

Direct: The rest of the difference, the difference of *c'* units, is due to the effect of *X* on *Y* independent of *M*.



Direct effect = c' = 0.062
Indirect effect = ab = -0.416(-0.289) = 0.120
Total effect = c = 0.062 + 0.120 = 0.182

Specific

Total: Two kids who differ by one scale point in ethnic harassment are estimated to differ by 0.182 units in perceived academic failure one year later, with the more frequently-harassed kid perceiving greater failure.

Indirect: They differ by 0.120 units in perceived failure as a result of the negative effect of harassment on self esteem a year later, which in turn increases perceived failure.

Direct: Independent of this mechanism, the more harassed kid is estimated to be 0.062 units higher in perceived failure.

It works for dichotomous *X* too.

Garcia, D. M., Schmitt, M. T., Branscombe, N. R., & Ellemers, N. (2010). Women's reactions to ingroup members who protest discriminatory treatment: The importance of beliefs about inequality and response appropriateness. ***European Journal of Social Psychology, 49***, 733-745.

European Journal of Social Psychology
Eur. J. Soc. Psychol. **40**, 733–745 (2010)
Published online 6 July 2009 in Wiley InterScience
(www.interscience.wiley.com) **DOI:** 10.1002/ejsp.644

Research article

Women's reactions to ingroup members who protest discriminatory treatment: The importance of beliefs about inequality and response appropriateness

DONNA M. GARCIA[1]*, MICHAEL T. SCHMITT[2], NYLA R. BRANSCOMBE[3] AND NAOMI ELLEMERS[4]
[1] *University of Guelph, Canada*
[2] *Simon Fraser University, Canada*
[3] *University of Kansas, USA*
[4] *Leiden University, The Netherlands*

Abstract

Our goal was to identify factors that shape women's responses to ingroup members who protest gender discrimination. We predicted and found that women who perceived gender discrimination as pervasive regarded a protest response as being more appropriate than a no protest response and expressed greater liking and less anger towards a female lawyer who protested rather than did not protest an unfair promotion decision. Further, beliefs about the appropriateness of the response to discrimination contributed to evaluations of the protesting lawyer. Perceptions that the complaint was an appropriate response to the promotion decision led to more positive evaluations of an ingroup discrimination protester. Copyright © 2009 John Wiley & Sons, Ltd.

Protest can be an effective means of improving the plight of a devalued group. Historically, there are many examples of protest, even from a single individual, that have advanced a group's social position (e.g. Meritor Savings Bank vs. Vinson, 477 US 57, 1986; Dekker vs. VJV-Centrum ECJ, 1992). Despite the potential gains to be obtained by protesting illegitimate treatment, protestors might not always be appreciated by members of their own group. Whether disadvantaged group members respond positively or negatively to ingroup protestors will likely depend upon the *perceived* implications that the protestor's action has for the ingroup. Unless protest is seen as justified by the social circumstances and an effective means of bringing about positive change, a protestor might be seen as making the ingroup look like complainers. Such threat to the ingroup's reputation could evoke the ire and disdain of the disadvantaged group towards the protestor. Hence, perceptions of the justification for and likely consequences of protest will be critical to others' reactions to an ingroup discrimination claimer. We propose that protest by an ingroup member will be seen as appropriate and thus appreciated to the extent that observers perceive that their ingroup is targeted by pervasive discrimination.

SOCIAL CONSEQUENCES OF CLAIMING DISCRIMINATION

Gender discrimination continues to be widespread throughout Western employment settings (see Charles & Grusky, 2004). The continuation of gender discrimination has substantive negative implications for women's economic and

*Correspondence to: Donna M. Garcia, Department of Psychology, University of Guelph Ontario, Guelph, Ontario, N1G 2W1Canada. E-mail: donnagarcia3@gmail.com

Copyright © 2009 John Wiley & Sons, Ltd.

Received 19 October 2008
Accepted 7 April 2009

Participants (all female) read a narrative about a female attorney who lost a promotion at her firm to a much less qualified male through unequivocally discriminatory actions of the senior partners.

Participants assigned to the 'protest' condition were then told she protested the decision by presenting an argument to the partners about how unfair the decision was.

Participants assigned to the 'no protest' condition were told that although she was disappointed, she accepted the decision and continued working at the firm.

After reading the narrative, the participants evaluated **how appropriate they perceived her response to be**, and also evaluated the characteristics of the attorney, the responses of which were aggregated to produce a measure of **"liking."** Prior to the study, the participants filled out the Modern Sexism Scale.

The data: PROTEST

*protest.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

	subnum	cond	sexism	angry	liking	respappr	protest
1	209	2	4.87	2	4.83	4.25	1.00
2	44	0	4.25	1	4.50	5.75	.00
3	124	2	5.00	3	5.50	4.75	1.00
4	232	2	5.50	1	5.66	7.00	1.00
5	30	2	5.62	1	6.16	6.75	1.00
6	140	1	5.75	1	6.00	5.50	1.00
7	27	2	5.12	2	4.66	5.00	1.00
8	64	0	6.62	1	6.50	6.25	.00
9	67	0	5.75	6	1.00	3.00	.00
10	182	0	4.62	1	6.83	5.75	.00
11	85	2	4.75	2	5.00	5.25	1.00
12	109	2	6.12	5	5.66	7.00	1.00
13	122	0	4.87	2	5.83	4.50	.00
14	69	1	5.87	1	6.50	6.25	1.00

SAS users, run this program to make a temporary or "work" data file named PROTEST.

protest

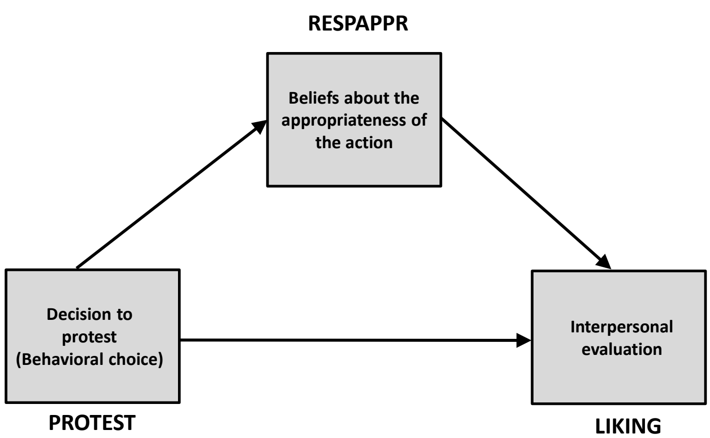
```
data protest;
input subnum cond sexism angry liking respappr protest;
cards;
209 2   4.87   2   4.83   4.25   1.00
44  0   4.25   1   4.50   5.75    .00
124 2   5.00   3   5.50   4.75   1.00
232 2   5.50   1   5.66   7.00   1.00
30  2   5.62   1   6.16   6.75   1.00
140 1   5.75   1   6.00   5.50   1.00
27  2   5.12   2   4.66   5.00   1.00
64  0   6.62   1   6.50   6.25    .00
67  0   5.75   6   1.00   3.00    .00
182 0   4.62   1   6.83   5.75    .00
85  2   4.75   2   5.00   5.25   1.00
109 2   6.12   5   5.66   7.00   1.00
122 0   4.87   2   5.83   4.50    .00
69  1   5.87   1   6.50   6.25   1.00
```

PROTEST: Experimental condition (1 = protest, 0 = no protest)

LIKING : Evaluation (liking) of the lawyer (higher = more positive evaluation, i.e. like more)

RESPAPPR: A measure of how appropriate the lawyer's behavior in response to the action of the partners was perceived to be for the situation (higher = more appropriate)

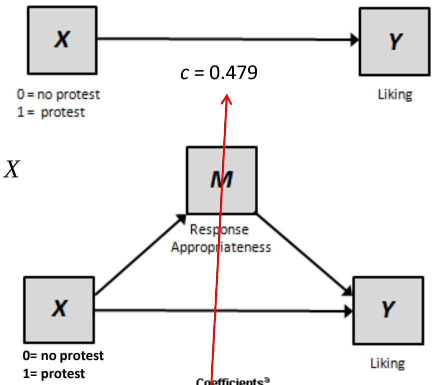
Our question



Do perceptions of the appropriateness of the response act as the mechanism through which that choice influences interpersonal evaluation?

Notice that this question is not asked contingent on evidence of simple association between the choice and the evaluation.

Using a set of OLS regression analyses



$$\hat{Y} = 5.313 + \underline{0.479}\,X$$

Because the two groups differ by one unit on *X*, this is the difference between the group means. The attorney was liked more when she protested than when she did not.

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	5.313	.161		33.055	.000
	X: experimental condition (0 = no protest, 1 = protest)	.479	.195	.213	2.460	.015

a. Dependent Variable: Y: liking of the target

SPSS:
```
regression/dep=liking/method=enter protest.
```

SAS:
```
proc reg data=protest;model liking=protest;run;
```

Interpretation when *X* is dichotomous

$$\boxed{\hat{Y}_j = 5.310 + 0.479\,X_j}$$

```
means tables = liking by protest.
```

Report

LIKING: liking of the target

PROTEST: experimental condition (0 = no protest, 1 = protest)	Mean	N	Std. Deviation
no protest	5.3102	41	1.30158
protest	5.7889	88	.87669
Total	5.6367	129	1.04970

When *X* = 1 (protest), $\hat{Y}$ = 5.310 + 0.479(1) = 5.789
When *X* = 0 (no protest), $\hat{Y}$ = 5.310 + 0.479(0) = 5.310

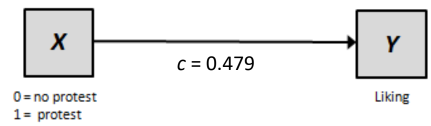
Notice that with X coded 0 and 1, the model yields the group means, *b* is the difference between the group means, and *the regression constant* is the mean for the group coded *X* = 0 (no protest condition).

More generally, if the two groups are coded by a difference of λ units, such that $X = \theta + \lambda$ for group 1 and X = θ for group 2,

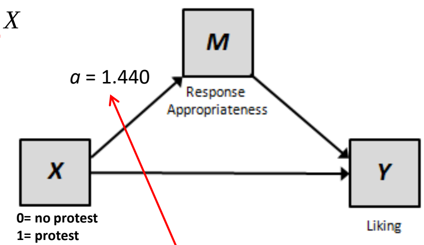
$$b = (\bar{Y}_1 - \bar{Y}_2)/\lambda$$

If you get in the habit of coding a dichotomous variable such that the groups differ by one unit on *X*, *b* will always be a difference between group means.

Using a set of OLS regression analyses



$$\hat{M} = 3.884 + \underline{1.440}\,X$$



Because the two groups differ by one unit on X, this is the difference between the group means. Protesting was seen as more appropriate for the situation than doing nothing.

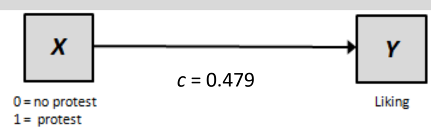
Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.884	.183		21.208	.000
	X: experimental condition (0 = no protest, 1 = protest)	$\boxed{1.440}$	.222	.499	6.493	.000

a. Dependent Variable: M: appropriateness of response

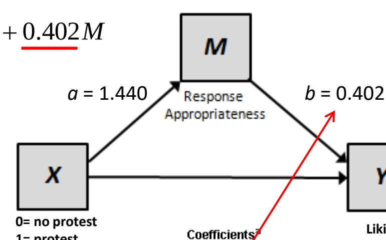
```
regression/dep=respappr/method=enter protest.
```

```
proc reg data=protest;model respappr=protest;run;
```

Using a set of OLS regression analyses



$$\hat{Y} = 3.751 - 0.100\,X + \underline{0.402}\,M$$



Holding constant what the attorney did, she was liked more by those who saw her behavior as more appropriate for the situation.

Coefficients[a]

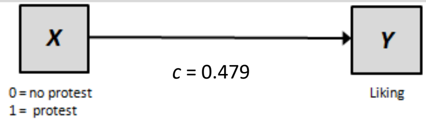
Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.751	.306		12.271	.000
	M: appropriateness of response	$\boxed{.402}$	.069	.517	5.789	.000
	X: experimental condition (0 = no protest, 1 = protest)	-.100	.200	-.045	-.501	.617

a. Dependent Variable: Y: liking of the target

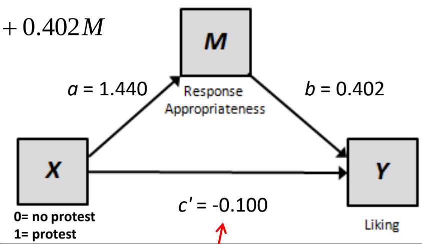
```
regression/dep=liking/method=enter respappr protest.
```

```
proc reg data=protest;model liking=respappr protest;run;
```

Using a set of OLS regression analyses



$$\hat{Y} = 3.751 - \underline{0.100}\,X + 0.402\,M$$



Because the two groups differ by one unit on *X*, this is the difference between the group means adjusted for differences between the groups in how appropriate her behavior was perceived as being for the situation (i.e., holding it constant)

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.751	.306		12.271	.000
	M: appropriateness of response	.402	.069	.517	5.789	.000
	X: experimental condition (0 = no protest, 1 = protest)	$\boxed{-.100}$	.200	-.045	-.501	.617

a. Dependent Variable: Y: liking of the target

```
regression/dep=liking/method=enter respappr protest.
```

```
proc reg data=protest;model liking=respappr protest;run;
```

Interpretation of total, direct, and indirect effects

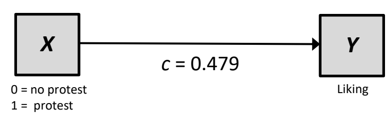
Generic

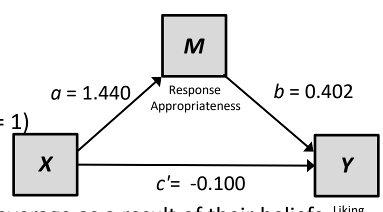
Total: Two people who differ by one unit on *X* are estimated to differ by *c* units on *Y* on average.

Indirect: They differ by *ab* units on average as a result of the effect of *X* on *M* which in turn affects *Y*.

Direct: The rest of the difference, the difference of *c′* units, is due to the effect of *X* on *Y* independent of *M*.

Direct effect = -0.100
Indirect Effect = 1.440(0.402) = 0.579
Total effect = -0.100 + 0.579 = 0.479





Specific

Total: Participants told the lawyer protested (*X* = 1) liked the lawyer 0.479 units **more**, on average, than those who were told she did not protest.

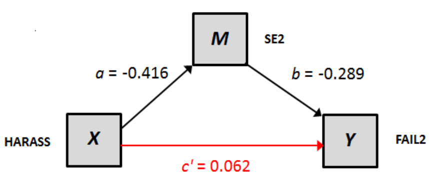
Indirect: They liked her by 0.579 units **more** on average as a result of their beliefs about the appropriateness of her response, which in turn affected their liking.

Direct: Among those equal in their beliefs about the appropriateness of her response, those who were told the lawyer protested liker her 0.100 units **less** (because the sign is negative) than those who were told she did not protest the decision.

Statistical Inference

Statistical inference: The direct effect

Inference for the direct effect is simple and noncontroversial. The inference can be framed in terms of a hypothesis test or a confidence interval. Any OLS regression program will provide both.



No statistically significant evidence of a direct effect of ethnic harassment on academic failure, $c' = 0.062$, $p = 0.396$, 95% CI = (-0.081 to 0.204)

```
regression/statistics defaults ci/dep=fail2/method=enter harass se2.
```

```
proc reg data=harass;model fail2=harass se2/stb clb;run;
```

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	2.385	.204		11.676	.000	1.983	2.786
	harass	.062	.072	.046	.850	.396	-.081	.204
	se2	-.289	.048	-.324	-5.988	.000	-.384	-.194

a. Dependent Variable: fail2

Statistical inference: The indirect effect

The indirect effect estimates the influence of *X* on *Y* through the mechanism represented by *M* (i.e., the $X \rightarrow M \rightarrow Y$ sequence). 21st-century mediation analysis bases claims of mediation on evidence that the indirect effect is different from zero.

A popular "20th-century" approach to inference: The Sobel test

Indirect effect

"Second order" estimator of the standard error of *ab*

$$Z = \frac{ab}{\sqrt{b^2 s_a^2 + a^2 s_b^2 + s_a^2 s_b^2}}$$

One version eliminates this term ("first order" estimator) ,

A *p*-value is derived by assuming normality of the sampling distribution of the indirect effect and using the standard normal distribution for derivation of *p*. A *p*-value no greater than α leads to the claim that the indirect effect is statistically different from zero at the α level of significance.

Computation with Harass Data

a = -0.416 s_a = 0.080

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.597	.123		29.123	.000
	harass	-.416	.080	-.276	-5.209	.000

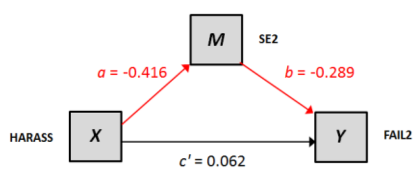
a. Dependent Variable: se2

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.385	.204		11.676	.000
	harass	.062	.072	.046	.850	.396
	se2	-.289	.048	-.324	-5.988	.000

a. Dependent Variable: fail2

b = -0.289 s_b = 0.048



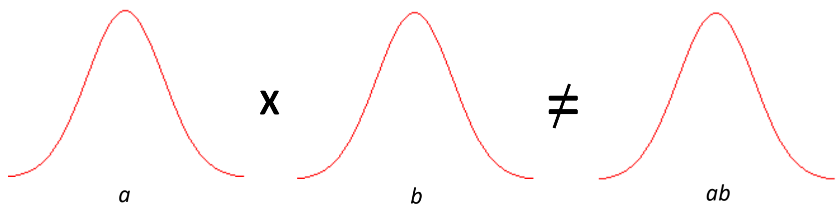
$$Z = \frac{ab}{\sqrt{b^2 s_a^2 + a^2 s_b^2 + s_a^2 s_b^2}}$$

$$Z = \frac{-0.416(-0.289)}{\sqrt{(-0.289)^2 0.080^2 + (-0.416)^2 0.048^2 + 0.080^2 0.048^2}} = \frac{0.120}{0.031} = 3.899,\ p = .0001$$

The indirect effect is statistically significant. But this test has serious problems.

What's wrong with the Sobel test?

For the Sobel test, the *p*-value is derived by assuming normality of the sampling distribution of the indirect effect and using the standard normal distribution. **Although this assumption is fairly sensible in large samples, it is not in smaller ones. What is a sufficiently large sample is situationally-specific, and typically you won't know going into the analysis whether or not to trust large sample theory.**



This assumption, which typically will not hold, yields a test that is lower in power than alternatives. **Experts in mediation analysis don't recommend the use of this test, though it remains popular.** Eventually, researchers will get the message.

The bootstrap confidence interval

Bootstrapping allows us to empirically estimate the sampling distribution of the indirect effect and generate a confidence interval (CI) for estimation and hypothesis testing. **It has become the preferred inferential method for estimating and testing indirect effects.**

(1) Take a random sample of size *n* from the sample *with replacement.*

(2) Estimate the indirect effect in this "resample".

(3) Repeat (1) and (2) a total of *k* times, where *k* is at least 1,000. The larger *k*, the better. I recommend at least 5,000.

(4) Use distribution of the indirect effect over multiple resamples as an approximation of the sampling distribution of the indirect effect.

(5) For 95% CI using "percentile" method, lower and upper bounds are 2.5th and 97.5th percentile in *k* bootstrap estimates of the indirect effect. Variations exist (e.g., 'bias corrected' or 'bias-corrected and accelerated' confidence intervals but they do not perform as well as percentile.)

Bootstrapping

Your data

X	M	Y
4.3	1.4	9.1
1.4	5.4	6.4
4.9	4.3	1.3
5.9	2.3	5.4
6.1	3.3	3.9
3.8	3.1	6.3
2.8	3.2	1.5
9.4	4.1	2.3
4.3	1.3	4.4
4.9	3.7	2.1

a = -0.051 b = -0.844

ab = 0.043

A resampling of your data

X	M	Y
5.9	2.3	5.4
4.9	4.3	1.3
9.4	4.1	2.3
4.9	4.3	1.3
4.3	1.4	9.1
1.4	5.4	6.4
3.8	3.1	6.3
9.4	4.1	2.3
6.1	3.3	3.9
4.3	1.3	4.4

a = 0.020 b = -0.921

ab = -0.018

Bootstrapping

Your data

X	M	Y
4.3	1.4	9.1
1.4	5.4	6.4
4.9	4.3	1.3
5.9	2.3	5.4
6.1	3.3	3.9
3.8	3.1	6.3
2.8	3.2	1.5
9.4	4.1	2.3
4.3	1.3	4.4
4.9	3.7	2.1

a = -0.051 b = -0.844

ab = 0.043

Another resampling of your data

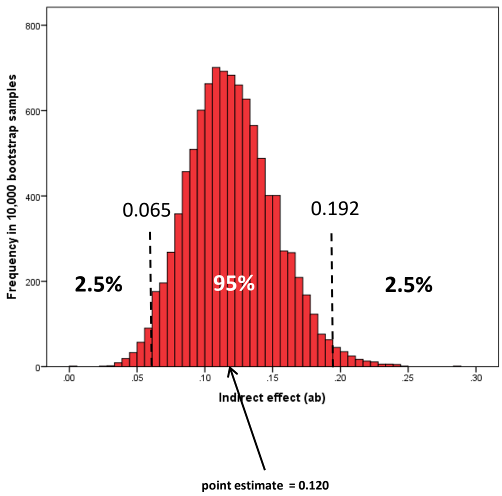
X	M	Y
6.1	3.3	3.9
4.9	4.3	1.3
2.8	3.2	1.5
4.9	3.7	2.1
3.8	3.1	6.3
9.4	4.1	2.3
2.8	3.2	1.5
4.9	4.3	1.3
4.9	3.7	2.1
1.4	5.4	6.4

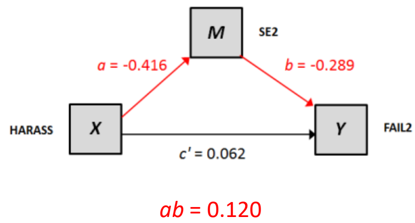
a = -0.034 b = 0.523

ab = -0.017

10,000 bootstrap estimates of the indirect effect

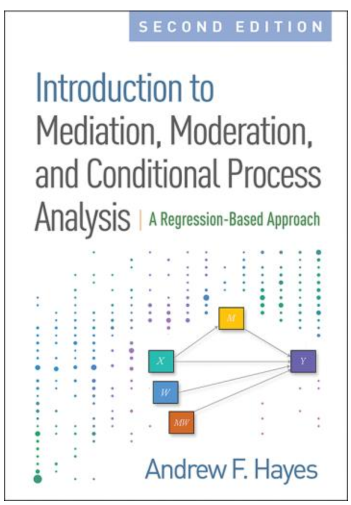
95% of the 10,000 bootstrap estimates of the indirect effect were between 0.065 and 0.192. This is our 95% confidence interval.





Zero is not in the confidence interval, so we **can** claim an indirect effect different from zero with 95% confidence. This is akin to (though not exactly the same as) rejecting the null hypothesis of no indirect effect at the α = 0.05 level of significance.

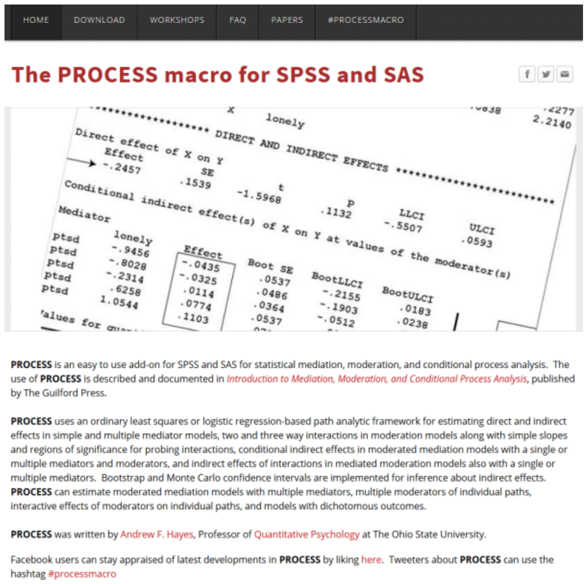
PROCESS



Published in December 2017 and available through The Guilford Press, Amazon.com, and elsewhere.

- First released in beta form in March of 2012 and later documented in Hayes (2013, IMMCPA, published by The Guilford Press).
- Available for both SPSS (in macro and "custom dialog" form) and SAS.
- An integration of functions available in my other published macros for mediation and moderation analysis (SOBEL, INDIRECT, MODMED, MODPROBE, MED3C) and a whole lot more, all in one command.
- A handy tool for both "confirmatory" and "exploratory" approaches to data analysis.
- Freely available at www.processmacro.org. The current release is v3.

processmacro.org



HOME | DOWNLOAD | WORKSHOPS | FAQ | PAPERS | #PROCESSMACRO

The PROCESS macro for SPSS and SAS

PROCESS is an easy to use add-on for SPSS and SAS for statistical mediation, moderation, and conditional process analysis. The use of **PROCESS** is described and documented in *Introduction to Mediation, Moderation, and Conditional Process Analysis*, published by The Guilford Press.

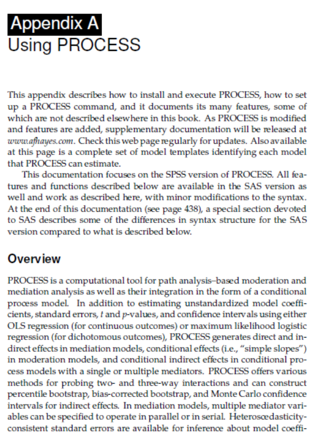
PROCESS uses an ordinary least squares or logistic regression-based path analytic framework for estimating direct and indirect effects in simple and multiple mediator models, two and three way interactions in moderation models along with simple slopes and regions of significance for probing interactions, conditional indirect effects in moderated mediation models with a single or multiple mediators and moderators, and indirect effects of interactions in mediated moderation models also with a single or multiple mediators. Bootstrap and Monte Carlo confidence intervals are implemented for inference about indirect effects. **PROCESS** can estimate moderated mediation models with multiple mediators, multiple moderators of individual paths, interactive effects of moderators on individual paths, and models with dichotomous outcomes.

PROCESS was written by Andrew F. Hayes, Professor of Quantitative Psychology at The Ohio State University.

Facebook users can stay appraised of latest developments in **PROCESS** by liking here. Tweeters about **PROCESS** can use the hashtag #processmacro

Read the documentation (eventually)

The PROCESS documentation is an eventual must-read. It describes how to use PROCESS, as well as its various options, capabilities, and limitations. It is available as Appendix A in Hayes (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis (2nd Ed)*. At a minimum, you must have the model templates handy, as PROCESS expects you to tell it which model number you are estimating and which variables play what role. You have the templates PDF.



Appendix A
Using PROCESS

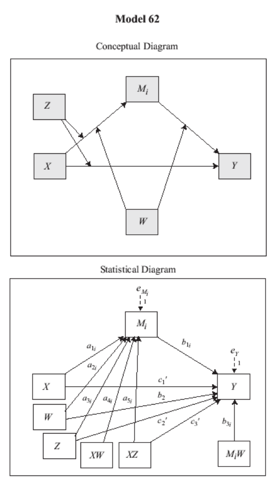
This appendix describes how to install and execute PROCESS, how to set up a PROCESS command, and it documents its many features, some of which are not described elsewhere in this book. As PROCESS is modified and features are added, supplementary documentation will be released at *www.afhayes.com*. Check this web page regularly for updates. Also available at this page is a complete set of model templates identifying each model that PROCESS can estimate.

This documentation focuses on the SPSS version of PROCESS. All features and functions described below are available in the SAS version as well and work as described here, with minor modifications to the syntax. At the end of this documentation (see page 438), a special section devoted to SAS describes some of the differences in syntax structure for the SAS version compared to what is described below.

Overview

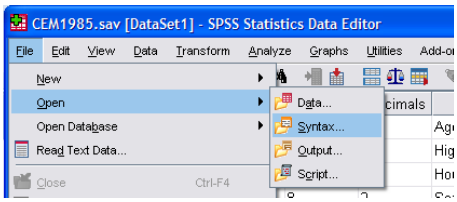
PROCESS is a computational tool for path analysis–based moderation and mediation analysis as well as their integration in the form of a conditional process model. In addition to estimating unstandardized model coefficients, standard errors, *t* and *p*-values, and confidence intervals using either OLS regression (for continuous outcomes) or maximum likelihood logistic regression (for dichotomous outcomes), PROCESS generates direct and indirect effects in mediation models, conditional effects (i.e., "simple slopes") in moderation models, and conditional indirect effects in conditional process models with a single or multiple mediators. PROCESS offers various methods for probing two- and three-way interactions and can construct percentile bootstrap, bias-corrected bootstrap, and Monte Carlo confidence intervals for indirect effects. In mediation models, multiple mediator variables can be specified to operate in parallel or in serial. Heteroscedasticity-consistent standard errors are available for inference about model coeffi-

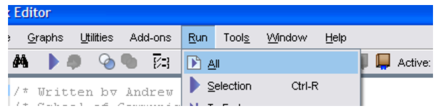
419



PROCESS as a syntax-driven macro

Open process.sps as a syntax file and run the entire program **exactly as is.** This produces a new SPSS command called PROCESS. See the documentation for details on the syntax structure. PROCESS goes away when you close SPSS.



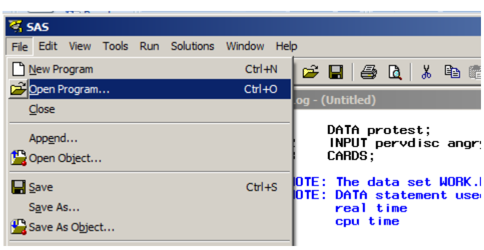


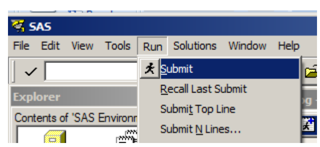
Once PROCESS is defined, you can run a properly-formatted PROCESS command in a new syntax window. Such a command might look something like below. Not all of the arguments below are required.

```
process y=react/x=cond/m=pmi/total=1/normal=1
        /model=4/boot=10000.
```

PROCESS for SAS

In SAS, open process.sas and submit the entire program **exactly as is.** This produces a new SAS command called %PROCESS. The syntax structure is described in the documentation.





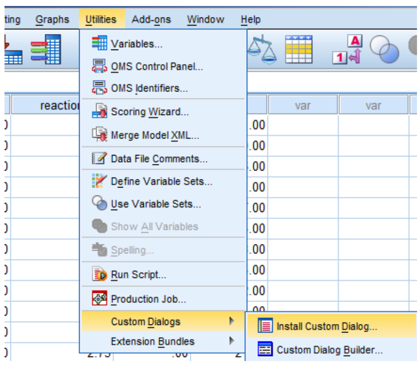
Once PROCESS is defined, you can run a properly-formatted PROCESS command in a new syntax window. Such a command might look something like below. Not all of the arguments below are required.

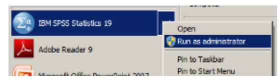
```
%process (data=pmi,y=reaction,x=cond,m=pmi,total=1,normal=1,
                  boot=10000, model=4);
```

PROCESS “Custom Dialog”

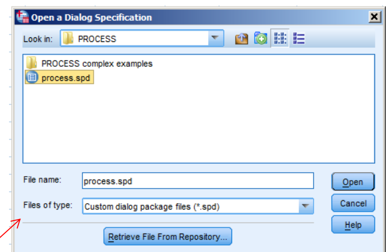
The PROCESS macro must be run at least once in your SPSS session to activate the PROCESS command. Custom Dialog files are permanently installed in SPSS, integrating the procedure into SPSS menus. Use the procedure below. In Windows, installation requires administrative access to your machine. You **probably** have to open SPSS as an administrator as well. You may not have access to do so.

In SPSS 23 and earlier:





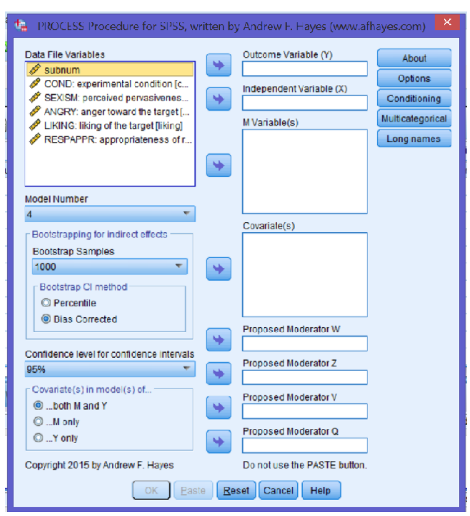
Right click SPSS to select



In SPSS24, look under “Extensions” for the Utilities option

When installed, PROCESS can be found under “Analyze”→”Regression”

PROCESS “Custom Dialog”

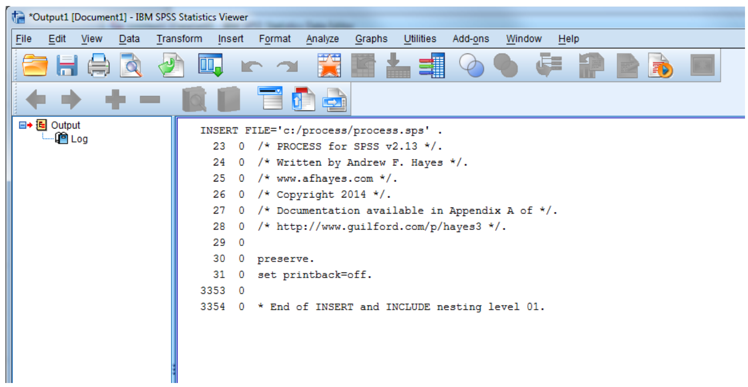


Installing the dialog box does not eliminate the need to run the PROCESS code if you plan on executing with syntax. And don’t use the PASTE button.

Autoexecution

It is possible to get SPSS to execute the PROCESS code on its own when SAS/SPSS executes. A document is provided to you with the course files that provides instructions for SPSS for Windows. I do not know how to do this with SPSS for Mac.

When successful, SPSS for Windows users typically see something like the following in the output window when SPSS is opened:



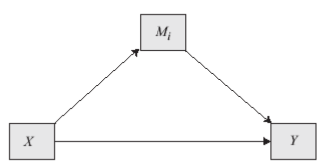
Model template system

PROCESS has 56 model templates, referred to by number in PROCESS, which specify the form of the model (moderation, mediation, moderated mediation, and so forth) and which variables play what roles.

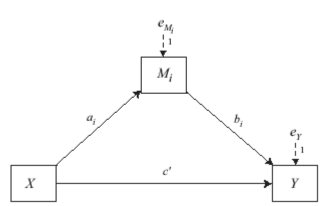
Example #1:

Model 4 is a simple or parallel multiple mediator model, which estimates the direct and indirect effect(s) of *X* on *Y* through one or more mediators (*M*) (up to 10 mediators at once)

Conceptual diagram



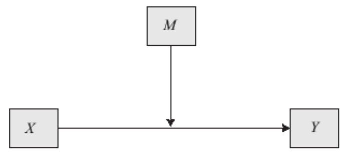
Statistical diagram



Model template system

PROCESS has 56 model templates, referred to by number in PROCESS, which specify the form of the model (moderation, mediation, moderated mediation, and so forth) and which variables play what roles.

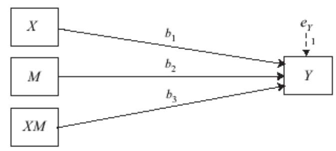
Conceptual diagram



Example #2:

Model 1 is a simple moderation model, with *M* moderating the effect of *X* on *Y*.

Statistical diagram

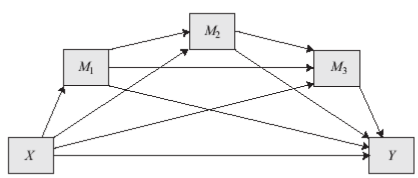


The statistical diagram shows the model in the form of a path diagram. This is the form in which the model is estimated.

Model template system

PROCESS has 56 model templates, referred to by number in PROCESS, which specify the form of the model (moderation, mediation, moderated mediation, and so forth) and which variables play what roles.

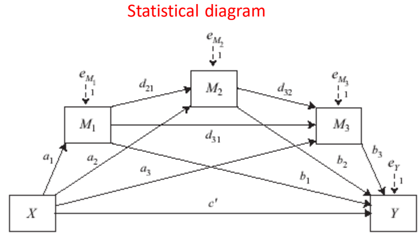
Conceptual diagram



Example #3:

Model 6 is a serial multiple mediator model, which estimates the direct and indirect effect(s) of *X* on *Y* through up to 4 mediators (*M*) chained together in serial. An example with **three** mediators is depicted to the right.

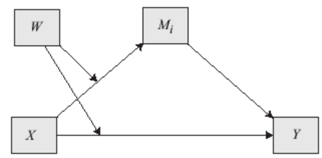
Statistical diagram



Model template system

PROCESS has 56 model templates, referred to by number in PROCESS, which specify the form of the model (moderation, mediation, moderated mediation, and so forth) and which variables play what roles.

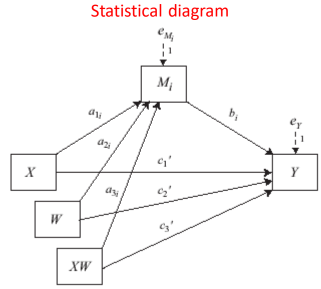
Conceptual diagram



Example #4:

Model 8 is a conditional process model which estimates the conditional direct and indirect effects of *X* on *Y* through *M*, with direct effect and "first stage" moderation by *W*.

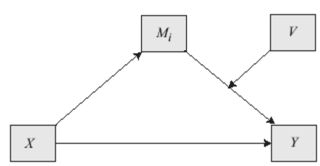
Statistical diagram



Model template system

PROCESS has 56 model templates, referred to by number in PROCESS, which specify the form of the model (moderation, mediation, moderated mediation, and so forth) and which variables play what roles.

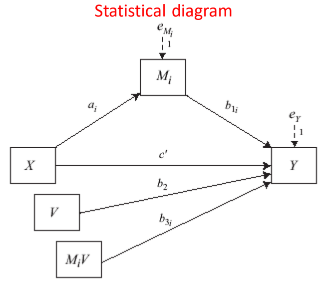
Conceptual diagram



Example #5:

Model 14 is a conditional process model which estimates the direct effect of *X* on *Y* and conditional indirect effects of *X* on *Y* through *M*, with "second stage" moderation by *V*.

Statistical diagram



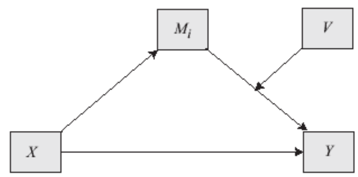
Model template system

PROCESS has 56 model templates, referred to by number in PROCESS, which specify the form of the model (moderation, mediation, moderated mediation, and so forth) and which variables play what roles.

Minimum required specifications

- Which variables play which role in the model (**y=** **x=** **m=** and so forth)
- Model number (**model=**)
- SAS only: Data file (**data=**)

Conceptual diagram



SPSS

```
PROCESS /y=yvar/x=xvar/m=mvlist/v=vvar/model=14.
```

SAS

```
%process (data=filename,y=yvar,x=xvar,m=mvlist,v=vvar,model=14);
```

Limitations and constraints

- Only one *X* and one *Y* allowed in a model.
- PROCESS is an OLS modeling tool. Categorical mediators or *Y* are not allowed.
- Up to 10 mediators in numbered models, 6 in custom models.
- No more than two moderators can be used in any model.
- A variable can play only one role in the model. For example a variable can't be both a moderator and a covariate, or both a mediator and a moderator.
- PROCESS is a single-level observed variable modeling system. No multilevel problems can be analyzed with PROCESS.
- PROCESS requires complete data. Listwise deletion is used for cases missing on any variable in the model.
- Although PROCESS will accept them, it is safer to restrict variable names to eight characters or fewer.

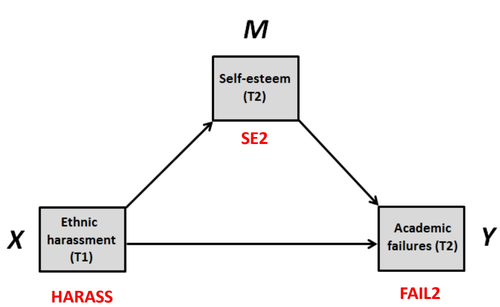
If you are familiar with PROCESS v2, see the "What's new in PROCESS 3" pages in your course book.

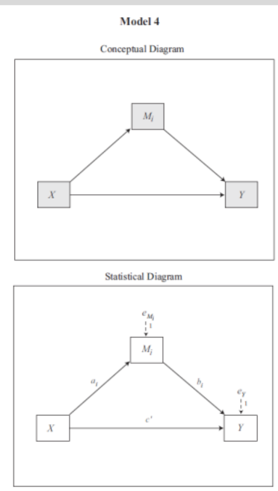
Differences between V2 & V3

- No longer need the `vars` list
- Covariates now listed in `cov` list
- Moderators are always `W` and `Z`, no more `V`, `M` or `Q` moderators
- No dichotomous Y
- A variety of models have been cut, but new ability to create and edit models
- New models for serial moderated mediation and serial and parallel mediation
- Probing option now defaults to what used to be `quantiles`, can use `moments` argument for legacy output
- Probing and plotting for models with any moderation
- Default is now percentile bootstrap, no more BC or ABC
- Multicategorical X or Moderators
- `Wmodval` and `zmodval` allow for multiple values
- Covariate assignment
- Bootstrap Cis for regression coefficients
- Model construction
- `Cluster`, `ws`, `varorder`, and `percent` are no longer options

Estimation of the harassment model in PROCESS

PROCESS Model 4






```
Process y=fail2/x=harass/m=se2/model=4/boot=10000
        /normal=1/percent=1/total=1.
```

```
%process (data=harass,y=fail2,x=harass,m=se2,model=4,
          boot=10000,normal=1,percent=1,total=1);
```

PROCESS output

```
*************** PROCESS Procedure for SPSS Version 3.00 ****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 4
    Y  : fail2
    X  : harass
    M  : se2

Sample
Size:  330

**************************************************************************
OUTCOME VARIABLE:
 se2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .2764      .0764      .2905    27.1349     1.0000   328.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.5966      .1235    29.1227      .0000     3.3536     3.8395
harass       -.4156      .0798    -5.2091      .0000     -.5725     -.2586
```

Output C

$$\hat{M} = 3.597 - 0.416X$$

path *a*

PROCESS output

```
**************************************************************************
Outcome: fail2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .3397      .1154      .2215    21.3247     2.0000   327.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     2.3845      .2042    11.6757      .0000     1.9827     2.7863
se2          -.2887      .0482    -5.9879      .0000     -.3836     -.1939
harass        .0616      .0725      .8499      .3960     -.0810      .2042

************************ TOTAL EFFECT MODEL ****************************
Outcome: fail2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .1356      .0184      .2451     6.1419     1.0000   328.0000      .0137

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.3460      .1134    11.8660      .0000     1.1229     1.5692
harass        .1816      .0733     2.4783      .0137      .0374      .3257

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

$$\hat{Y} = 2.385 + 0.062X - 0.289M$$

path *b*

path *c'*

$$\hat{Y} = 1.346 + 0.182X$$

path *c*

Output C

PROCESS output

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

Total effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .1816      .0733     2.4783      .0137      .0374      .3257

Direct effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .0616      .0725      .8499      .3960     -.0810      .2042

Indirect effect of X on Y
        Effect    Boot SE   BootLLCI   BootULCI
se2      .1200      .0321      .0629      .1899

Normal theory tests for indirect effect
     Effect         se          Z          p
      .1200      .0308     3.8993      .0001
```

path *c*

path *c*′

ab with 95% bootstrap confidence interval

Sobel test

Kids one unit higher in harassment frequency were 0.416 units lower in self esteem one year later (a = -0.416), and lower self-esteem was related to higher perceived academic failure (*b* = -0.289). So harassment indirectly affected perceived academic failure (point estimate: 0.120, 95% bootstrap CI = 0.063 to 0.190). After accounting for this mechanism, there was no evidence of an effect of harassment on perceived academic failure (direct effect = 0.062, *p* = 0.396, 95% CI = -0.081 to 0.204)

Output C

Some additional options

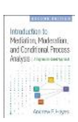
SPSS

```
process y=fail2/x=harass/m=se2/model=4/boot=10000/normal=1/total=1/
         effsize=1/conf=99/save=1/seed=25545.
```

SAS

```
%process (data=harass, y=fail2,x=harass,m=se2,model=4,boot=10000,
         normal=1,total=1,effsize=1,conf=99,save=boots,seed=25545);
```

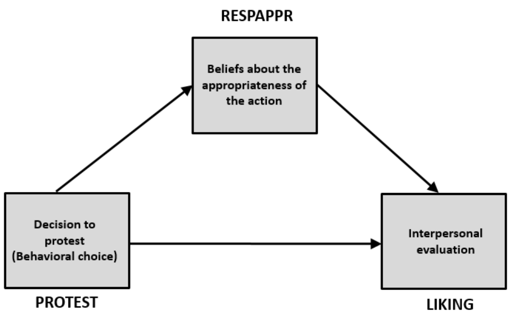
EFFSIZE=1: Generates various effect size measures for the indirect effect.
CONF=z: Changes level of confidence to z% for confidence intervals.
SAVE=1 or SAVE=fn: in SPSS, produces a file of all bootstrap estimates of all regression coefficients in the model. In SAS, saves bootstrap estimates to a file named "fn"
SEED=xxxx: Seeds the random number generator for replication of resamples over repeated runs of PROCESS.

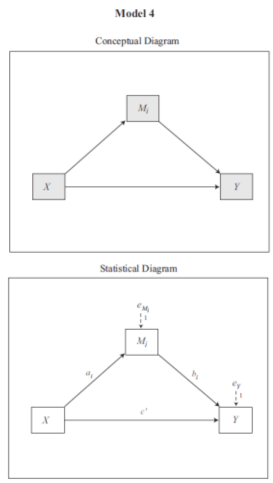


See the documentation in Appendix A of IMMCPA for details.

Estimation of the PROTEST model in PROCESS

PROCESS Model 4



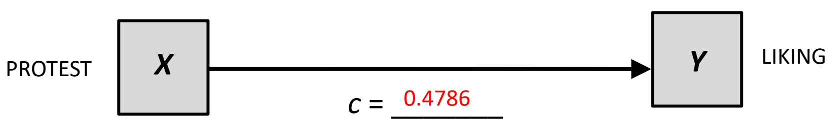


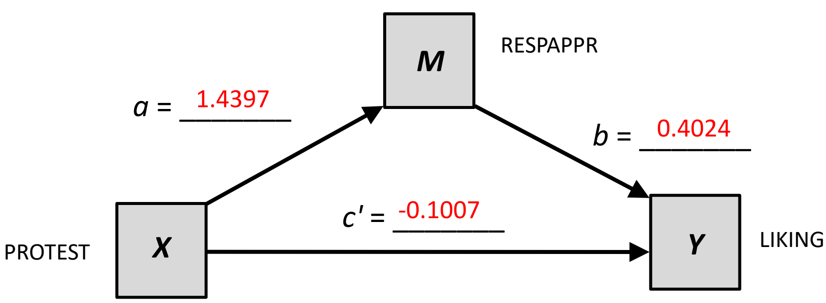
What should the PROCESS command look like?

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

not required

```
%process (data=protest, y=liking,x=protest,m=respappr,model=4,boot=10000,normal=1,total=1);
```





Indirect effect = 0.5793, 95% bootstrap CI = .3113 to .9067

PROCESS output

```
*************** PROCESS Procedure for SPSS Version 3.00 ****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 4
    Y  : liking
    X  : protest
    M  : respappr

Sample
Size:  129

**************************************************************************
OUTCOME VARIABLE:
 respappr

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .4992      .2492     1.3753    42.1550     1.0000   127.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.8841      .1831    21.2078      .0000     3.5217     4.2466
protest      1.4397      .2217     6.4927      .0000     1.0009     1.8785
**************************************************************************
```

$$\hat{M} = 3.884 + 1.440X$$

path *a*

PROCESS output

```
**************************************************************************
Outcome: liking

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .4959      .2459      .8441    20.5483     2.0000   126.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.7473      .3058    12.2553      .0000     3.1422     4.3524
respappr      .4024      .0695     5.7884      .0000      .2648      .5400
protest      -.1007      .2005     -.5023      .6163     -.4975      .2960

************************* TOTAL EFFECT MODEL ****************************
Outcome: liking

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .2131      .0454     1.0601     6.0439     1.0000   127.0000      .0153

Model
              coeff         se          t          p       LLCI       ULCI
constant     5.3102      .1608    33.0244      .0000     4.9921     5.6284
protest       .4786      .1947     2.4584      .0153      .0934      .8639
**************************************************************************
```

$$\hat{Y} = 3.747 - 0.101X + 0.402M$$

path *b*

path *c'*

$$\hat{Y} = 5.310 + 0.479X$$

path *c*

PROCESS output

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

Total effect of X on Y
     Effect         SE          t          p        LLCI        ULCI
      .4786      .1947     2.4584      .0153       .0934       .8639     path c

Direct effect of X on Y
     Effect         SE          t          p        LLCI        ULCI
     -.1007      .2005     -.5023      .6163      -.4975       .2960     path c′

Indirect effect of X on Y
             Effect    Boot SE   BootLLCI   BootULCI
respappr      .5793      .1519      .3113      .9067     ab with 95% bootstrap
                                                         confidence interval
Normal theory tests for indirect effect
     Effect         se          Z          p
      .5793      .1350     4.2924      .0000          Sobel test
```

Her behavior was perceived as more appropriate if she protested relative to when she did not (*a* = 1.440), and the more appropriate her behavior, the more positively she was perceived (*b* = 0.402). Her choice to protest had a positive effect on how favorably she was perceived indirectly through perceived appropriateness of the response (point estimate: 0.579, 95% CI = 0.311 to 0.907). After accounting for this mechanism, there was no effect of her choice to protest on how she was evaluated (direct effect = -0.101, *p* = 0.62)

Confounding

Kids who reported greater harassment earlier reported lower in self-esteem later, but they also reported lower self-esteem earlier (r = -0.18), and self-esteem was temporally consistent over time (r = 0.51). Furthermore, students who reported lower self esteem later reported greater academic failure later, but these later low self-esteem students also reported higher greater academic failure earlier (r = -0.26), and academic failure was temporally consistent (r = 0.30)

Correlations

		harass	se1	fail1	se2	fail2
harass	Pearson Correlation	1	-.176	.196	-.276	.136
	Sig. (2-tailed)		.001	.000	.000	.014
	N	330	330	330	330	330
se1	Pearson Correlation	-.176	1	-.306	.505	-.259
	Sig. (2-tailed)	.001		.000	.000	.000
	N	330	330	330	330	330
fail1	Pearson Correlation	.196	-.306	1	-.255	.297
	Sig. (2-tailed)	.000	.000		.000	.000
	N	330	330	330	330	330
se2	Pearson Correlation	-.276	.505	-.255	1	-.337
	Sig. (2-tailed)	.000	.000	.000		.000
	N	330	330	330	330	330
fail2	Pearson Correlation	.136	-.259	.297	-.337	1
	Sig. (2-tailed)	.014	.000	.000	.000	
	N	330	330	330	330	330

Pre-existing self-esteem and failure confound the relationships we believe to be causal. We want to know whether later self-esteem and academic failure are related to ethnic harassment frequency after accounting for initial self-esteem and failure.

Confounding

Some effects in a mediation model are subject to 'confounding' <u>even when *X* is based on random assignment</u>, making causality harder to establish. Partialing out various confounders can help though won't solve the problem entirely.

$$\hat{Y} = i_1 + cX + c_2U$$
$$\hat{M} = i_2 + aX + a_2U$$
$$\hat{Y} = i_3 + c'X + bM + c_2'U$$

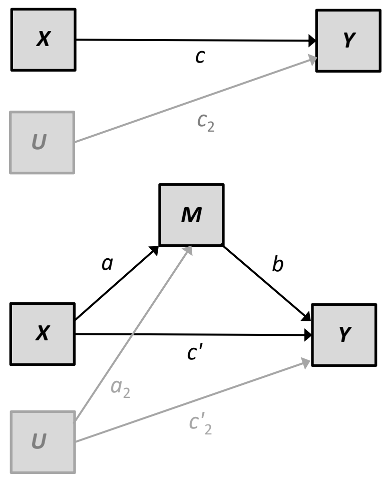
total effect = direct effect + indirect effect

$c \quad = \quad c' \quad + \quad (a \times b)$

indirect effect = total effect – direct effect

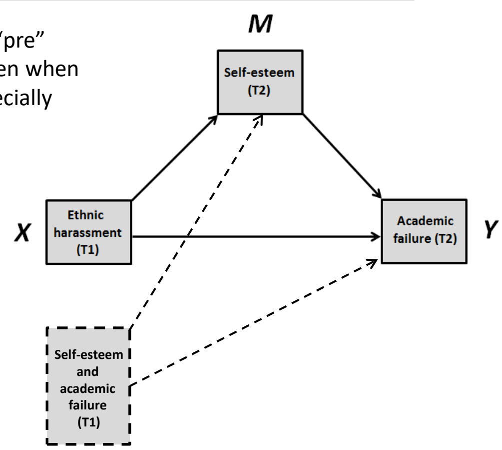
$a \times b \quad = \quad c \quad - \quad c'$

A simple mediation model, adjusting for a potential confounding variable (*U*)



Some rationales for adjusting for prior state

When available, it is desirable to include "pre" measures of *M* and/or *Y* as covariates, even when *X* is experimentally manipulated, but especially when it is not.

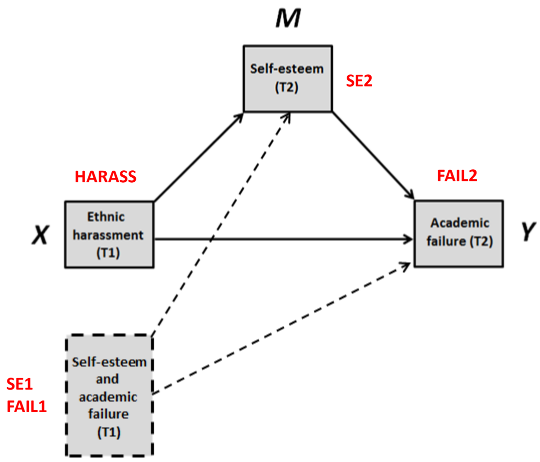


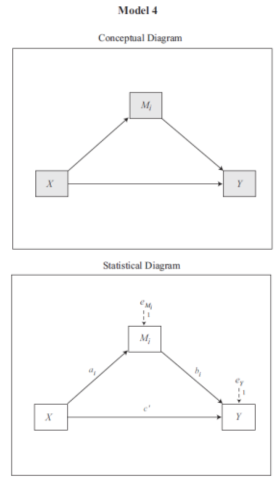
(a) Doing so can increase precision in the estimation of *X*'s effect on *M* and/or *Y* if pre-measures are correlated with later measures (as they typically are).

(b) Prior states often are correlated with *X*, *M*, or *Y*, introducing a "self-selection" threat to causal claims. Including prior state helps to reduce that threat.

(c) It gives an interpretation to paths that are closer to a "change" interpretation without regression artifacts that can be introduced with the use of difference scores. In this example, the *b* path estimates the relationship between later self-esteem and how much higher or lower a student's failure is given expected later failure from prior self-esteem and failure. Path *a* estimates the relationship between ethnic harassment frequency and how much lower or higher a student's self-esteem is later relative to what would be expected given his or her earlier self-esteem and academic failure.

Adding covariates to a model using PROCESS






```
process cov=se1 fail1/y=fail2/x=harass/m=se2/model=4/
  boot=10000/total=1.
```

```
%process (data=harass, cov=se1 fail1,y=fail2,x=harass,m=se2,
  model=4,boot=10000,total=1);
```

PROCESS output

Output D

```
*************** PROCESS Procedure for SPSS Version 3.00 ****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 4
    Y  : fail2
    X  : harass
    M  : se2

Covariates:
 se1       fail1

Sample
Size:  330

**************************************************************************
OUTCOME VARIABLE:
 se2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5450      .2971      .2224    45.9259     3.0000   326.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     2.0280      .2412     8.4095      .0000     1.5536     2.5024
harass       -.2728      .0717    -3.8025      .0002     -.4139     -.1317
se1           .4879      .0536     9.1081      .0000      .3825      .5933
fail1        -.1010      .0606    -1.6661      .0967     -.2202      .0182

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

path *a*

PROCESS output

Output D

```
**************************************************************************
Outcome: fail2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .4059      .1648      .2105    16.0276     4.0000   325.0000      .0000

Model
```

	coeff	se	t	p	LLCI	ULCI	
constant	2.0934	.2588	8.0900	.0000	1.5843	2.6025	
se2	-.2175	.0539	-4.0375	.0001	-.3235	-.1115	path *b*
harass	.0196	.0713	.2754	.7832	-.1207	.1599	path *c′*
se1	-.0672	.0584	-1.1517	.2503	-.1820	.0476	
fail1	.2307	.0592	3.8966	.0001	.1142	.3471	

```
************************* TOTAL EFFECT MODEL ***************************
Outcome: fail2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .3505      .1229      .2203    15.2219     3.0000   326.0000      .0000

Model
```

	coeff	se	t	p	LLCI	ULCI	
constant	1.6523	.2400	6.8842	.0000	1.1801	2.1245	
harass	.0790	.0714	1.1060	.2695	-.0615	.2194	path *c*
se1	-.1733	.0533	-3.2513	.0013	-.2782	-.0685	
fail1	.2526	.0603	4.1886	.0000	.1340	.3713	

PROCESS output

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

Output D

Total effect of X on Y

Effect	SE	t	p	LLCI	ULCI	
.0790	.0714	1.1060	.2695	-.0615	.2194	path *c*

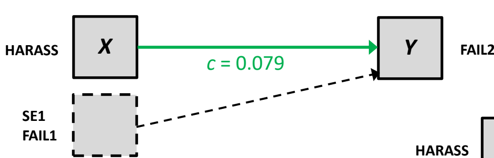
Direct effect of X on Y

Effect	SE	t	p	LLCI	ULCI	
.0196	.0713	.2754	.7832	-.1207	.1599	path *c′*

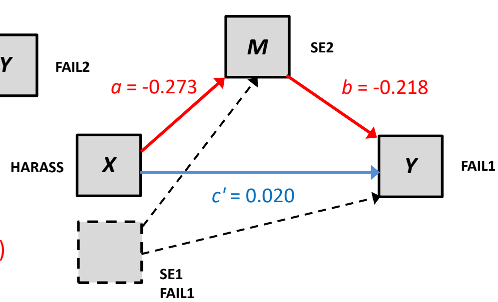
Indirect effect of X on Y

	Effect	Boot SE	BootLLCI	BootULCI
se2	.0593	.0229	.0203	.1089

ab with 95% bootstrap confidence interval



$$c = c' + a \times b$$
$$0.079 = 0.020 + (-0.273)(-0.218)$$
$$0.079 = 0.020 + 0.059$$



What about Baron & Kenny?

Also called the "causal steps" approach, it was popularized by Baron and Kenny (1986) as a test of mediation.

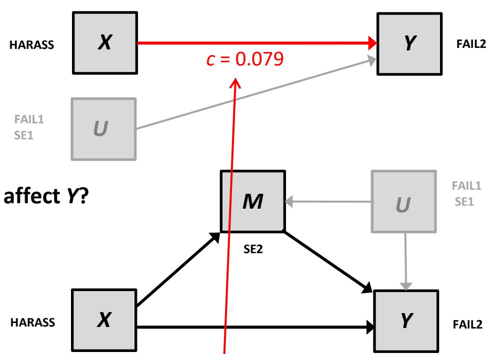
Conditions required to claim *M* functions as a mediator of the relationship between *X* and *Y*:

(1) Does *X* affect *Y?*
(2) Does *X* affect *M* ?
(3) Does *M* affect *Y* holding *X* constant ?

..as gauged by a hypothesis test.

(4) Is the direct effect of *X* closer to zero than the total effect?
  (i) if direct effect is closer to zero than total effect but statistically different from zero, claim "partial mediation"
  (ii) If direct effect is closer to zero than total effect and not statistically different from zero, claim "complete mediation"

Using a set of OLS regression analyses



Condition 1: Does *X* affect *Y*?

NO. THUS, NO MEDIATION BY THE CAUSAL STEPS STRATEGY

Coefficients[a]

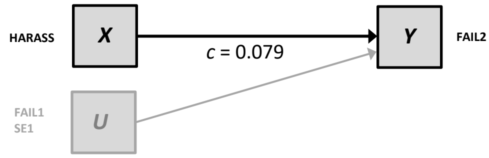
Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.652	.240		6.884	.000
	harass	.079	.071	.059	1.106	.270
	fail1	.253	.060	.231	4.189	.000
	se1	-.173	.053	-.179	-3.251	.001

a. Dependent Variable: fail2

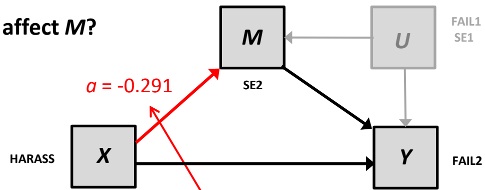
```
regression/dep=fail2/method=enter harass fail1 se1.
```

```
proc reg data=harass;model fail2=harass fail1 se1;run;
```

Using a set of OLS regression analyses



Condition 2: Does *X* affect *M*?



Coefficients[a]

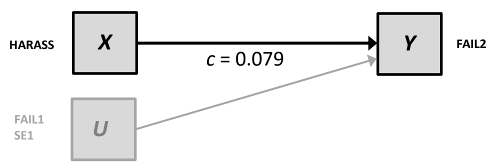
Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.824	.208		8.751	.000
	harass	-.291	.071	-.193	-4.091	.000
	se1	.513	.052	.471	9.951	.000

a. Dependent Variable: se2

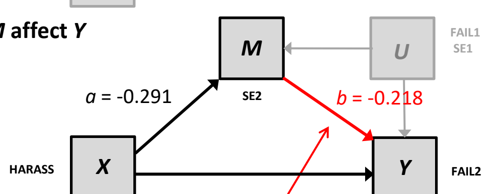
```
regression/dep=se2/method=enter harass se1.
```

```
proc reg data=harass;model se2=harass se1;run;
```

Using a set of OLS regression analyses



Condition 3: Does *M* affect *Y* holding *X* constant?



Coefficients[a]

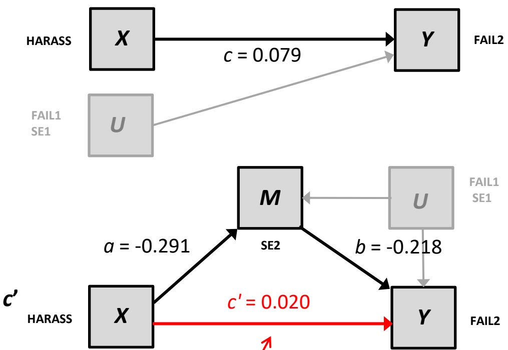
Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.093	.259		8.090	.000
	harass	.020	.071	.015	.275	.783
	se2	-.218	.054	-.244	-4.038	.000
	fail1	.231	.059	.211	3.897	.000
	se1	-.067	.058	-.069	-1.152	.250

a. Dependent Variable: fail2

```
regression/dep=fail2/method=enter harass se2 fail1 sel.
```

```
proc reg data=harass;model fail2=harass se2 fail1 sel;run;
```

Using a set of OLS regression analyses



Direct effect (c') is closer to zero than the total effect (*c*) and is not statistically different from zero. But can we really call this 'complete mediation' given that there was no total effect of *X* in the first place by commonly-used inferential rules?

4. **Qualitatively compare *c* to *c'***

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.093	.259		8.090	.000
	harass	.020	.071	.015	.275	.783
	se2	-.218	.054	-.244	-4.038	.000
	fail1	.231	.059	.211	3.897	.000
	se1	-.067	.058	-.069	-1.152	.250

a. Dependent Variable: fail2

```
regression/dep=fail2/method=enter harass se2 fail1 se1.
```

```
proc reg data=harass;model fail2=harass se2 fail1 se1;run;
```

Problems with the causal steps approach

- ❑ Indirect effect is logically inferred rather than directly estimated

  But typically, we make inferences from data using estimates of quantities pertinent to the question. Why should inferences about indirect effects be any different?

  A fallacious rebuttal: if *a* and *b* are both different from zero (as established by rejection of the null hypothesis) so too must their product, so no estimate or test of indirect effect is needed.

  a) Although frequently that will be true, it isn't necessarily true.
  b) An indirect effect may be different from zero even in the absence of evidence that both paths *a* and *b* are.

- ❑ If data fail to meet a single criterion, **game over--no indirect effect through *M*.**

  The use of multiple, fallible hypothesis tests gives this approach the **lowest power** among competing methods for testing intervening variable effects. **Tests or claims of mediation should not be based on the significance of individual paths in the model.**

Problems with the causal steps approach

- ❑ If total effect (path *c*) is not detectably different from zero, the **game doesn't even begin.**

  This is logically sensible if you accept one definition of a mediator variable – a variable that is causally between *X* and *Y* and that accounts for their association.

  - By this definition, an effect that does not exist can't be mediated. But the significance of *c* neither constrains nor determines the size of the product of paths *a* and *b*, nor does it tell us whether that product is different from zero.
  - Kenny and Judd (2014, *Psychological Science*) illustrate that a hypothesis test about the total effect is generally less powerful than a hypothesis test about the indirect effect.

- ❑ Because the indirect effect is not quantified, this method does not lend itself well to comparisons between indirect effects in multiple mediators models, or to modeling of the size of indirect effects ('conditional process analysis')

"Complete"/"full" and "partial" mediation

The causal steps strategy is often used as a means of labeling a process as "complete" or "partial mediation". There is little value to this semantic labeling exercise.

- ❑ What if there is no evidence of a total effect (i.e., *c* non-significant)? This can happen, and actually does more often than people probably realize. Thus, these concepts don't have a place much of the time.
- ❑ The reliance on statistical significance criteria means that when power is high for the test on *c*′ partial mediation is the best you can hope for, and when power is relatively low, complete mediation is more likely. So if establishing complete mediation is your goal, <u>you should intentionally limit the size of your sample to as small as necessary</u>
- ❑ Establishing complete mediation by your favored mediator does not preclude others from being able to make the same claim with their own favored mediator.
- ❑ "Direct effects" don't exist in reality. All effects are mediated by something. Thus, a claim of 'partial mediation' is a claim that one has not specified the model correctly.

Experts in mediation analysis are abandoning these concepts. They are of historical interest only these days.

Mediation analysis summary thus far

- Mediators are variables which are causally located between two variables *X* and *Y* and that explain, in part, the effect of *X* on *Y*. *X* affects *M* which in turn affects *Y*.

- The causal steps strategy popularized by Baron and Kenny (1986) remains a popular method for mediation analysis.
  - Yet it is among the lowest in power, in some circumstances, massively so.
  - It is not consistent with modern thinking about mediation analysis.
  - Its use is not recommended. Soon you won't be able to get away with it.

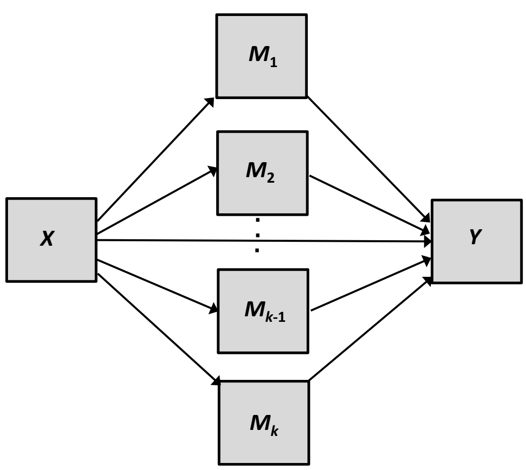
- Tests of mediation should be based on an estimate of the indirect effect.
  - Sobel test for inference in large samples only, but we don't know how large is large enough.
  - Bootstrap confidence intervals in a sample of any size.

- There is no need to condition the hunt for an indirect effect on a statistically significant total effect (path *c*).

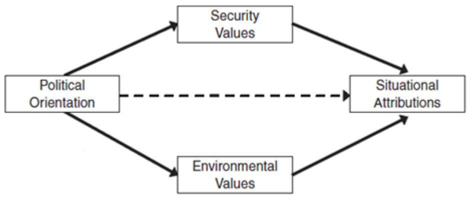
- Focus interpretation on the size and sign of the indirect effect. Tests of significance for the individual paths ($X \rightarrow M$ and $M \rightarrow Y$) are useful as supplemental information but need not be part of the story.

Models with More Than One Mediator

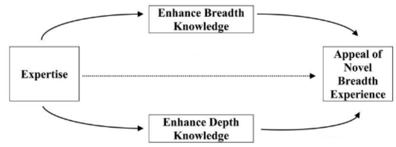
A parallel multiple mediator model



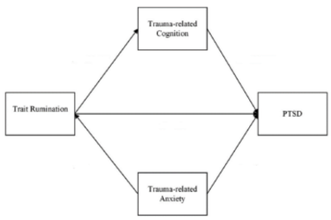
Some examples From the literature with 2 mediators



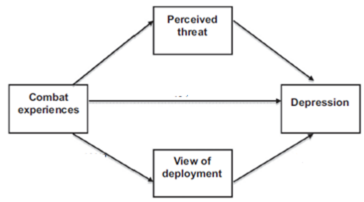
Morgan, G. S., Mullen, E., & Skitka, L. J. (2010). When values and attributions collide: Liberals' and conservatives' values motivate attributions for alleged misdeeds. ***Personality and Social Psychology Bulletin***, 36, 1241-1254.



Clarkson, J. J., Janiszewski, C., & Cinelli, M. D. (2013). The desire for consumption knowledge. ***Journal of Consumer Research, 39,*** 1313-1329.

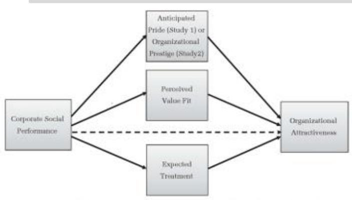


Spinhoven, P., Penninx, B. W., Krempeniou, A., et al. (2015). Trait rumination predicts onset of post-traumatic stress disorder through trauma-related cognitive appraisals: A 4-year longitudinal study. ***Behaviour Research and Therapy, 71***, 101-109.

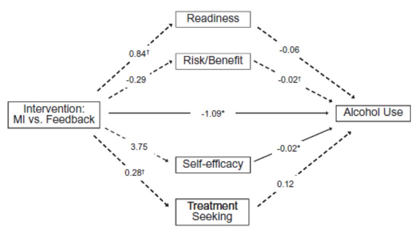


Pitts, B. L., & Safer, M. A. (2016). Retrospective appraisals mediate the effect of combat experiences on PTS and depression symptoms in U.S. Army medics**.** ***Journal of Traumatic Stress, 29***, 65-71.

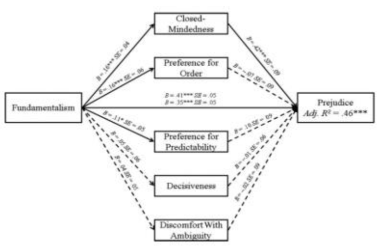
Some examples from the literature with several mediators



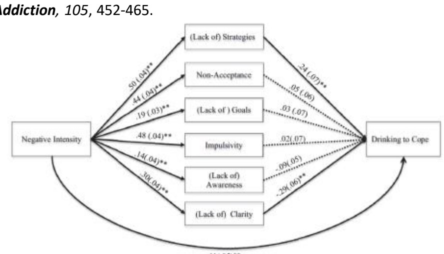
Jones, D. A., Willness, C. R., & Madey, S. (2014). Why are job seekers attracted by corporate social performance? Experimental and field tests of three signal-based mechanisms. ***Academy of Management Journal, 57***, 383-404.



Barnett, N. P., Apodaca, T. R., et al. (2010). Moderators and mediators of two brief interventions for alcohol in the emergency department. ***Addiction***, ***105***, 452-465.



Brandt, M. J., & Reyna, C. (2010). The role of prejudice and the need for closure in religious fundamentalism. ***Personality and Social Psychology Bulletin, 36***, 715-725.



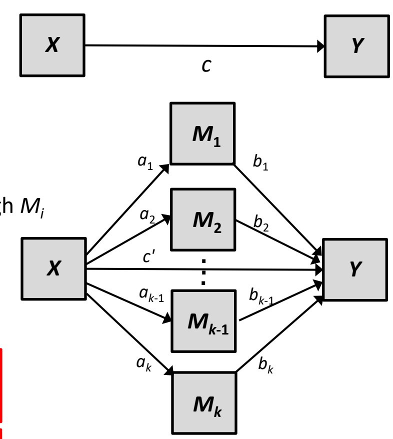
Veilleux, J. C.,Skinner, K. D., Reese, E. D., & Shaver, J. A. (2014). Negative affect intensity influences drinking to cope through facets of emotion dysregulation. ***Personality and Individual Differences***, ***49***, 96-101.

Why estimate such a model?

- Many causal effects probably operate through multiple mechanisms simultaneously. Better to estimate a model consistent with such real-world complexities.
- If your proposed mediator is correlated with the real mediator but not caused by the independent variable, a model with only your proposed mediator in it will be a misspecification and will potentially misattribute the process to your proposed mediator rather than the real mediator—"epiphenomenality."
- Different theories may postulate different mediators as mechanisms. Including them all in a model simultaneously allows for a formal statistical comparison of indirect effects representing different theoretical mechanisms.

Path Analysis: Total, Direct, and Indirect Effects

$$\hat{Y} = i_1 + cX$$
$$\hat{M}_j = i_j + a_j X$$
$$\hat{Y} = i_3 + c'X + \sum_{j=1}^{k} b_j M_j$$



c = "total effect" of X on Y

$a_j \times b_j$ = "specific indirect effect" of X on Y through M_i

$\Sigma(a_j \times b_j)$= "total indirect effect" of X on Y

c' = "direct effect" of X on Y

total effect = direct effect + total indirect effect

$c \quad = \quad c' \quad + \quad \Sigma(a_j \times b_j)$

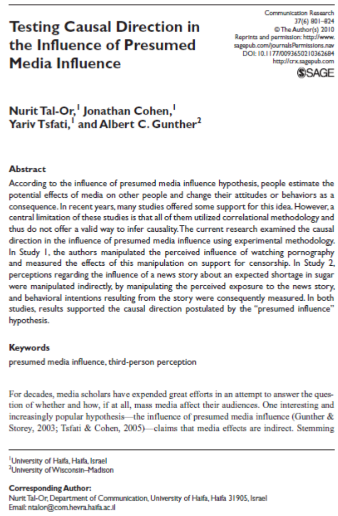
total indirect effect = total effect – direct effect

$\Sigma(a_j \times b_j) \quad = \quad c \quad - \quad c'$

Example: PMI

Participants read a news story describing global economic conditions that are leading to a possible sugar shortage and increase in prices. They were told either that the article was **about to be** published on the front page of Israel's largest daily ("**front page**" condition) or in the inside of an economic supplement ("**back page**" condition). Participants were randomly assigned to condition. This is the main independent variable.

Participants were asked a series of questions used to gauge their intention to buy sugar (soon, and how much). This was the main dependent variable (**reported reaction**). Higher = greater intention to buy. They were also asked a set of questions to assess (1) the extent to which publication of the article would influence the public to buy sugar (**presumed media influence**) and (2) how important the topic was to the general public (**perceived importance**). Higher = greater presumed influence/greater perceived importance.



Communication Research
37(6) 801–824
© The Author(s) 2010
Reprints and permission: http://www.sagepub.com/journalsPermissions.nav
DOI: 10.1177/0093650210362684
http://crx.sagepub.com
SAGE

Testing Causal Direction in the Influence of Presumed Media Influence

Nurit Tal-Or,[1] Jonathan Cohen,[1] Yariv Tsfati,[1] and Albert C. Gunther[2]

Abstract

According to the influence of presumed media influence hypothesis, people estimate the potential effects of media on other people and change their attitudes or behaviors as a consequence. In recent years, many studies offered some support for this idea. However, a central limitation of these studies is that all of them utilized correlational methodology and thus do not offer a valid way to infer causality. The current research examined the causal direction in the influence of presumed media influence using experimental methodology. In Study 1, the authors manipulated the perceived influence of watching pornography and measured the effects of this manipulation on support for censorship. In Study 2, perceptions regarding the influence of a news story about an expected shortage in sugar were manipulated indirectly, by manipulating the perceived exposure to the news story, and behavioral intentions resulting from the story were consequently measured. In both studies, results supported the causal direction postulated by the "presumed influence" hypothesis.

Keywords

presumed media influence, third-person perception

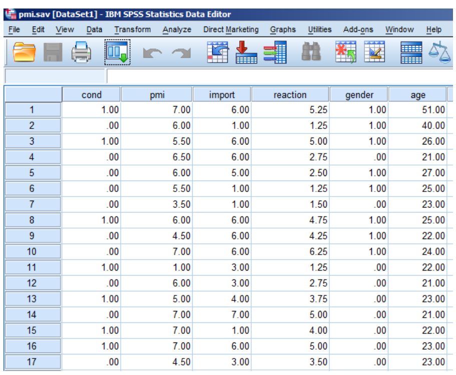
For decades, media scholars have expended great efforts in an attempt to answer the question of whether and how, if at all, mass media affect their audiences. One interesting and increasingly popular hypothesis—the influence of presumed media influence (Gunther & Storey, 2003; Tsfati & Cohen, 2005)—claims that media effects are indirect. Stemming

[1]University of Haifa, Haifa, Israel
[2]University of Wisconsin–Madison

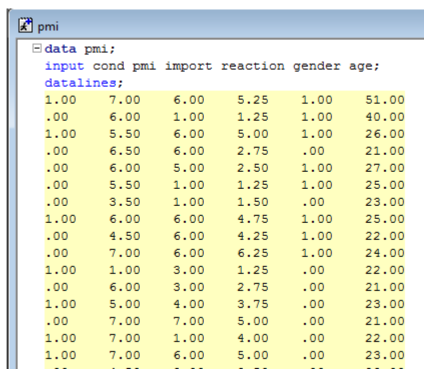
Corresponding Author:
Nurit Tal-Or, Department of Communication, University of Haifa, Haifa 31905, Israel
Email: ntalor@com.hevra.haifa.ac.il

Question: Does the location of the article affect behavioral intentions (i.e., their reported reaction) indirectly through presumed media influence, through perceived importance, or both? Would people told the article was to be published on the front page perceive it to be more important, and also infer greater influence on the public than those told it would be published in an economic supplement, which would in turn lead to greater intentions to act?

The data: PMI



	cond	pmi	import	reaction	gender	age
1	1.00	7.00	6.00	5.25	1.00	51.00
2	.00	6.00	1.00	1.25	1.00	40.00
3	1.00	5.50	6.00	5.00	1.00	26.00
4	.00	6.50	6.00	2.75	.00	21.00
5	.00	6.00	5.00	2.50	1.00	27.00
6	.00	5.50	1.00	1.25	1.00	25.00
7	.00	3.50	1.00	1.50	.00	23.00
8	1.00	6.00	6.00	4.75	1.00	25.00
9	.00	4.50	6.00	4.25	1.00	22.00
10	.00	7.00	6.00	6.25	1.00	24.00
11	1.00	1.00	3.00	1.25	.00	22.00
12	.00	6.00	3.00	2.75	.00	21.00
13	1.00	5.00	4.00	3.75	.00	23.00
14	.00	7.00	7.00	5.00	.00	21.00
15	1.00	7.00	1.00	4.00	.00	22.00
16	1.00	7.00	6.00	5.00	.00	23.00
17	.00	4.50	3.00	3.50	.00	23.00




```
data pmi;
input cond pmi import reaction gender age;
datalines;
1.00   7.00   6.00   5.25   1.00   51.00
.00    6.00   1.00   1.25   1.00   40.00
1.00   5.50   6.00   5.00   1.00   26.00
.00    6.50   6.00   2.75   .00    21.00
.00    6.00   5.00   2.50   1.00   27.00
.00    5.50   1.00   1.25   1.00   25.00
.00    3.50   1.00   1.50   .00    23.00
1.00   6.00   6.00   4.75   1.00   25.00
.00    4.50   6.00   4.25   1.00   22.00
.00    7.00   6.00   6.25   1.00   24.00
1.00   1.00   3.00   1.25   .00    22.00
.00    6.00   3.00   2.75   .00    21.00
1.00   5.00   4.00   3.75   .00    23.00
.00    7.00   7.00   5.00   .00    21.00
1.00   7.00   1.00   4.00   .00    22.00
1.00   7.00   6.00   5.00   .00    23.00
```

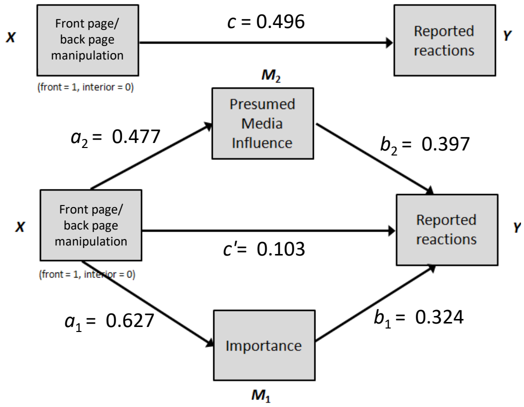
COND: Experimental condition (1 = front page, 0 = back page)

PMI : Presumed media influence (higher = greater presumed influence of media on public)

IMPORT: Perceived importance of sugar shortage (higher = more important)

REACTION: Response to article (higher = purchase sugar sooner and in higher quantity)

Example



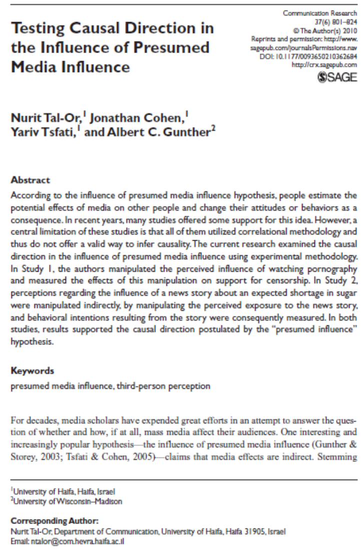
Direct effect = 0.103

Specific indirect effect via PMI: 0.477(0.397) = 0.190

Specific indirect effect via Import: 0.627(0.324) = 0.203

Total indirect effect = 0.190 + 0.203 = 0.393

Total effect = 0.190 + 0.203 + 0.103 = 0.496



Communication Research
37(6) 801–824

Reprints and permission: http://www.sagepub.com/journalsPermissions.nav
DOI: 10.1177/0093650210362684
http://crx.sagepub.com
SAGE

Testing Causal Direction in the Influence of Presumed Media Influence

Nurit Tal-Or,[1] Jonathan Cohen,[1] Yariv Tsfati,[1] and Albert C. Gunther[2]

Abstract

According to the influence of presumed media influence hypothesis, people estimate the potential effects of media on other people and change their attitudes or behaviors as a consequence. In recent years, many studies offered some support for this idea. However, a central limitation of these studies is that all of them utilized correlational methodology and thus do not offer a valid way to infer causality. The current research examined the causal direction in the influence of presumed media influence using experimental methodology. In Study 1, the authors manipulated the perceived influence of watching pornography and measured the effects of this manipulation on support for censorship. In Study 2, perceptions regarding the influence of a news story about an expected shortage in sugar were manipulated indirectly, by manipulating the perceived exposure to the news story, and behavioral intentions resulting from the story were consequently measured. In both studies, results supported the causal direction postulated by the "presumed influence" hypothesis.

Keywords

presumed media influence, third-person perception

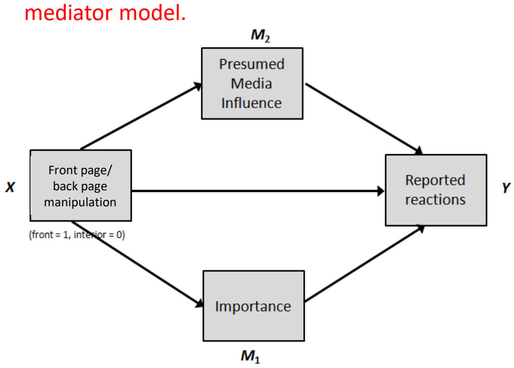
For decades, media scholars have expended great efforts in an attempt to answer the question of whether and how, if at all, mass media affect their audiences. One interesting and increasingly popular hypothesis—the influence of presumed media influence (Gunther & Storey, 2003; Tsfati & Cohen, 2005)—claims that media effects are indirect. Stemming

[1]University of Haifa, Haifa, Israel
[2]University of Wisconsin–Madison

Corresponding Author:
Nurit Tal-Or, Department of Communication, University of Haifa, Haifa 31905, Israel
Email: ntalor@com.hevra.haifa.ac.il

Estimation and inference using PROCESS

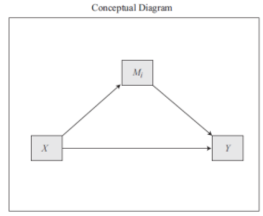
PROCESS model 4 is used for the parallel multiple mediator model.



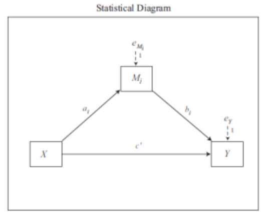
Up to 10 mediators can be listed in the "m =" list. Order does not matter.

Model 4

Conceptual Diagram



Statistical Diagram

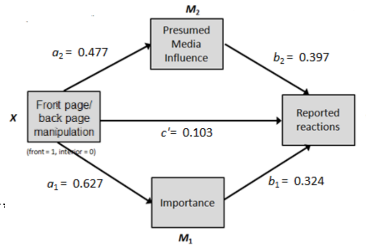


```
process y=react/x=cond/m=import pmi/total=1/boot=10000/model=4/normal=1/contrast=1.
```

```
%process (data=pmi,y=reaction,x=cond,m=import pmi,total=1,boot=10000,model=4,
         normal=1,contrast=1);
```

PROCESS output

Output E




```
Model = 4
    Y = reaction
    X = cond
   M1 = import
   M2 = pmi

Sample size
        123

**************************************************************************
Outcome: import

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .1809      .0327     2.9411     4.0942     1.0000   121.0000      .0452

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.9077      .2127    18.3704      .0000     3.4866     4.3288
cond          .6268      .3098     2.0234      .0452      .0135     1.2401
**************************************************************************
Outcome: pmi

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .1808      .0327     1.7026     4.0878     1.0000   121.0000      .0454

Model
              coeff         se          t          p       LLCI       ULCI
constant     5.3769      .1618    33.2222      .0000     5.0565     5.6973
cond          .4765      .2357     2.0218      .0454      .0099      .9431
**************************************************************************
```

$$\hat{M}_1 = 3.098 + 0.627X$$

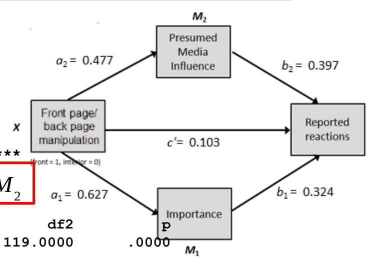
$$\hat{M}_2 = 5.377 + 0.477X$$

a_1 path

a_2 path

PROCESS output

Output E




```
**************************************************************************
Outcome: reaction

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5702      .3251     1.6628    19.1118     3.0000   119.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     -.1498      .5298     -.2828      .7778    -1.1989      .8993
import        .3244      .0707     4.5857      .0000      .1843      .4645
pmi           .3965      .0930     4.2645      .0000      .2124      .5806
cond          .1034      .2391      .4324      .6662     -.3701      .5768

************************** TOTAL EFFECT MODEL ****************************
Outcome: reaction

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .1603      .0257     2.3610     3.1897     1.0000   121.0000      .0766

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.2500      .1906    17.0525      .0000     2.8727     3.6273
cond          .4957      .2775     1.7860      .0766     -.0538     1.0452
```

$$\hat{Y} = -0.150 + 0.103X + 0.324M_1 + 0.397M_2$$

$$\hat{Y} = 3.250 + 0.496X$$

b_1 path

b_2 path

c' path

c path

PROCESS output

Output E

```
***************** TOTAL, DIRECT, AND INDIRECT EFFECTS ********************
Total effect of X on Y

     Effect         SE          t          p        LLCI       ULCI
      .4957      .2775     1.7860      .0766     -.0538     1.0452

Direct effect of X on Y

     Effect         SE          t          p        LLCI       ULCI
      .1034      .2391      .4324      .6662     -.3701      .5768

Indirect effect of X on Y

           Effect    Boot SE   BootLLCI   BootULCI
TOTAL       .3923      .1657      .0844      .7462
import      .2033      .1158      .0063      .4608
pmi         .1890      .1031      .0059      .4170
(C1)        .0144      .1435     -.2616      .3045

Normal theory tests for specific indirect effects

           Effect         se          Z          p
import      .2033      .1120     1.8154      .0695
pmi         .1890      .1057     1.7872      .0739

Specific indirect effect contrast definitions
(C1)   import     minus      pmi
```

c path

c' path

$a_1b_1 + a_2b_2$ with bootstrap CI

a_1b_1 with bootstrap CI

a_2b_2 with bootstrap CI

Sobel tests (less trustworthy than bootstrap CIs)

The data are consistent with the claim that article location influences reactions indirectly through both perceived importance (0.203; 95% CI = 0.006 to 0.461) and through presumed media influence (0.189; 95% CI = 0.005 to 0.417).

Things to consider

(1) In a multiple mediator model, the specific indirect effect through M_k quantifies the component of the total indirect effect that is unique to M_k.

M_k may function as a mediator variable when considered in isolation but not when considered with other mediator variables in the same model. If the intervening variables are highly intercorrelated, they can “cancel out” each others’ effects.

(2) It is possible for a total indirect effect to be not detectably different from zero even when one or more specific indirect effects is.

total indirect effect = sum of specific indirect effects

$$\Sigma (a_i b_i) \quad = \quad a_1b_1 + a_2b_2 + \ldots a_kb_k.$$

Scenario (a): A single large specific indirect combined with several tiny ones.

Scenario (b): Specific indirect effects that have different signs and add to near zero.

In multiple mediator models, the total indirect effect is rarely of much interest.

Comparing specific indirect effects

Indirect effects quantify how *Y* changes as *X* changes by one unit through a mediator. **They are free of the scale of measurement of the mediators**. So in multiple mediator models, indirect effects linking the same *X* to the same *Y* are directly comparable even If the mediators are measured on different scales. We can statistically compare them if so desired. No standardization or other arithmetic gymnastics is required.

Approach #1: Calculate the ratio of the difference between the indirect effect through M_i and the indirect effect through M_j to its standard error. Assuming a normally distributed sampling distribution of the difference, a *p*-value for the null hypothesis that the difference equals zero can be derived from the standard normal distribution.

$$Z = \frac{a_i b_i - a_j b_j}{se_{a_i b_i - a_j b_j}}$$

Approach #2: Bootstrap a confidence interval for the $a_i b_i - a_j b_j$ and ascertain whether 0 is in the confidence interval as a pseudo null hypothesis test that the difference is zero.

PROCESS can generate a bootstrap confidence interval for all possible pairwise comparisons between specific indirect effects

PROCESS output

```
process y=react/x=cond/m=import pmi/total=1/boot=10000/model=4/normal=1/contrast=1.
```

```
%process (data=pmi,y=reaction,x=cond,m=import pmi,total=1,boot=10000,model=4,
         normal=1,contrast=1);
```

```
Indirect effect of X on Y

         Effect    Boot SE   BootLLCI   BootULCI
TOTAL     .3923     .1657      .0844     .7462
import    .2033     .1158      .0063     .4608
pmi       .1890     .1031      .0059     .4170
(C1)      .0144     .1435     -.2616     .3045

Normal theory tests for specific indirect effects

         Effect        se          Z         p
import    .2033     .1120     1.8154     .0695
pmi       .1890     .1057     1.7872     .0739

Specific indirect effect contrast definitions
 (C1)    import      minus       pmi
```

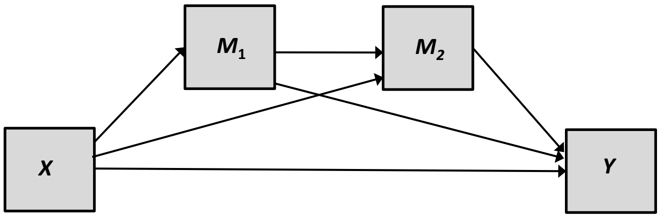
$a_1 b_1 - a_2 b_2$

Output E

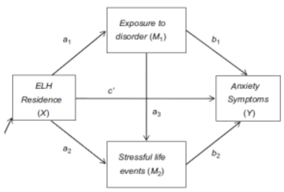
The specific indirect effect of article location on reactions through perceived importance is not statistically different from the specific indirect effect through presumed media influence (difference = 0.014; 95% CI = -0.262 to 0.305).

The serial multiple mediator model

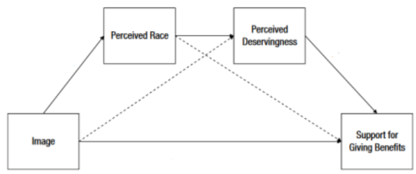
A serial multiple mediator model with two mediators and all possible direct and indirect effects freely estimated.



Some examples in the literature:

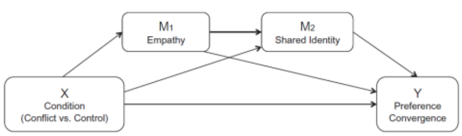


Casciano, R., & Massey, D. S. (2012). Neighborhood disorder and anxiety symptoms: New evidence from a quasiexperimental study. ***Health and Place,*** *18*, 180-190.

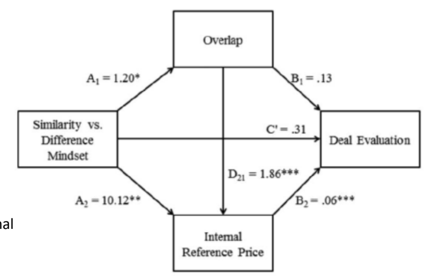


Brown-Iannuzzi, J. L., Dotsch, R., Cooley, E., & Payne, B. K. (2017). The relationship between mental representations of welfare recipients and attitudes toward welfare. ***Psychological Science, 28***, 92-103.

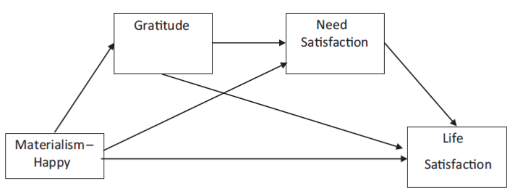
More examples from the literature



Schrift, R. Y., & Moty, A. (2015). Pain and preferences: Observed decisional conflict and the convergence of preferences. ***Journal of Consumer Research,*** *42*, 515-534..

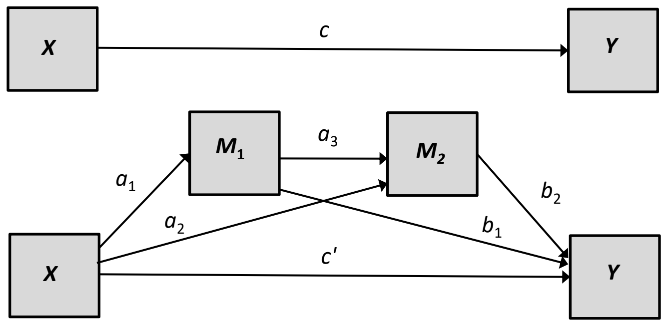


Kan, C., Lichtenstein, D. R., Grant, S. J., & Janiszwski, C. (2014). Strengthening the influence of advertised reference prices through information priming. ***Journal of Consumer Research,*** *40*, 1078-1096.



Tsang,J-A., Carpenter, T. P., Roberts, J. A., Frisch, M. B., & Carlisle, R. D. (2014). Why are materialists less happy? The role of gratitude and need satisfaction in the relationship between materialism and life satisfaction. ***Personality and Individual Differences,*** *64*, 62-66.

Serial mediation: Path analysis rules



The total effect of *X* on *Y* is equal to the direct effect of *X* plus the sum of all specific indirect effects (there are three of them here).

$$\hat{Y} = i_Y + cX$$
$$\hat{M}_1 = i_1 + a_1 X$$
$$\hat{M}_2 = i_2 + a_2 X + a_3 M_1$$
$$\hat{Y} = i_3 + c'X + b_1 M_1 + b_2 M_2$$

Direct effect of *X*: c'
Specific indirect effect of *X* through M_1: $a_1 b_1$
Specific indirect effect of *X* through M_2: $a_2 b_2$
Specific indirect effect of *X* through M_1 and M_2: $a_1 a_3 b_2$
Total indirect effect of *X*: $a_1 b_1 + a_2 b_2 + a_1 a_3 b_2$
Total effect of *X*: $c = c' + a_1 b_1 + a_2 b_2 + a_1 a_3 b_2$

Example

May 1st, 2011, 11:30PM

Professor Erik Nisbet (OSU School of Communication) had an national telephone survey in the field examining perceptions of Muslims in the U.S. when Obama announces the death of bin Laden; 390 respondents prior to announcement (BINLADEN = 0) and 271 after announcement (BINLADEN = 1). See report in materials provided for details.

Measures

STEREO: Stereotype endorsement, 4 items (5-pt semantic differential)

"Please tell us how much you associate each of the following sets of characteristics with Muslims"

e.g., Peaceful – Violent
Tolerant – Fanatical

RTHREAT: Realistic threat, 5 items (5-pt Likert)

"Below are a few statements expressing different views about Muslims living in the U.S. Please read and tell us how much you agree with each statement"

e.g., "Muslims in the U.S. sympathize with terrorists"
"Muslims make America a more dangerous place to live"

MCIVIL: Restriction of Muslim civil liberties, 5 items (5-pt Likert)

"Below are some statements people have expressed about Muslim civil liberties and terrorism in the U.S. Please read each and tell us how much you agree or disagree…"

e.g., "All Muslims in the U.S. should be required to carry a special ID card"
"Muslims in the U.S. should register their whereabouts with the U.S. government"

Higher reflect greater negative stereotype endorsement/threat/willingness to restrict…"

The Data: BINLADEN

BINLADEN.SAV

binladen.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

	binladen	rthreat	stereo	mcivil	age	ideo	sex
1	0	3.00	2.80	2.80	7.5		
2	1	2.00	1.80	3.20	3.3		
3	1	2.25	2.00	2.80	5.6		
4	1	2.00	2.60	3.40	4.0		
5	1	4.00	4.20	4.00	5.9		
6	0	1.00	1.40	4.80	7.9		
7	0	4.00	5.00	5.00	5.8		
8	0	4.00	2.80	3.80	7.2		
9	0	3.25	3.00	2.80	5.7		
10	1	4.00	5.00	5.00	5.9		
11	0	2.50	2.20	2.60	5.0		
12	1	1.50	3.20	1.60	3.9		
13	1	3.25	3.00	3.40	5.6		
14	1	4.25	2.80	5.00	4.6		
15	1	2.50	3.80	3.20	2.6		
16	1	2.75	3.40	3.20	5.6		

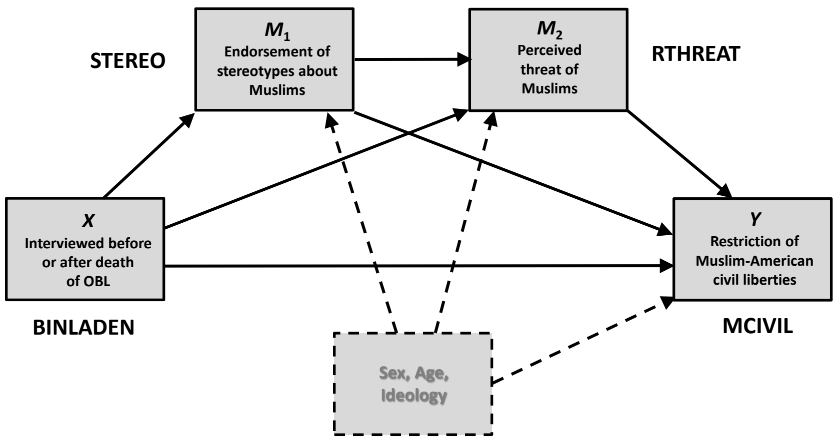
BINLADEN.SAS

binladen

```
data binladen;
input binladen rthreat stereo mcivil age ideo sex;
datalines;
0   3.00    2.80    2.80    7.5 8   0
1   2.00    1.80    3.20    3.3 4   1
1   2.25    2.00    2.80    5.6 6   0
1   2.00    2.60    3.40    4.0 5   1
1   4.00    4.20    4.00    5.9 8   1
0   1.00    1.40    4.80    7.9 6   0
0   4.00    5.00    5.00    5.8 5   0
0   4.00    2.80    3.80    7.2 6   1
0   3.25    3.00    2.80    5.7 4   1
1   4.00    5.00    5.00    5.9 5   1
0   2.50    2.20    2.60    5.0 7   1
1   1.50    3.20    1.60    3.9 4   1
1   3.25    3.00    3.40    5.6 9   0
1   4.25    2.80    5.00    4.6 4   1
```

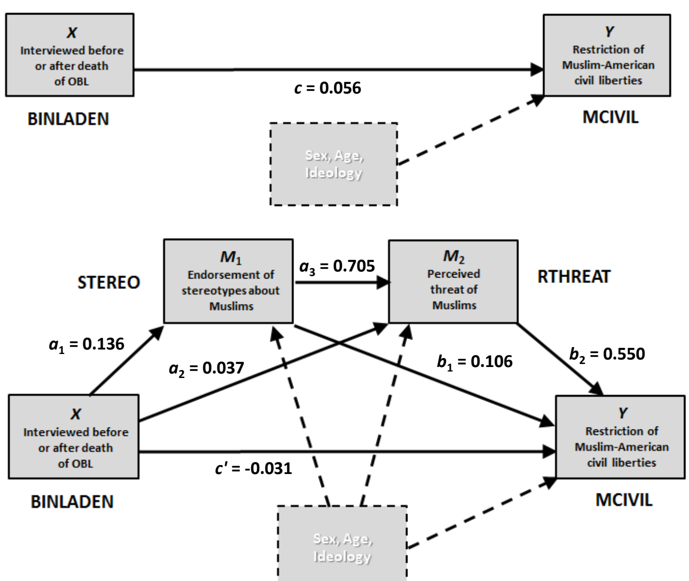
Also included in the data file are respondent **age** in <u>decades</u>, political **ideo**logy (7 point scale, higher = more conservative) and **gender** (0 = female, 1 =male).

Example

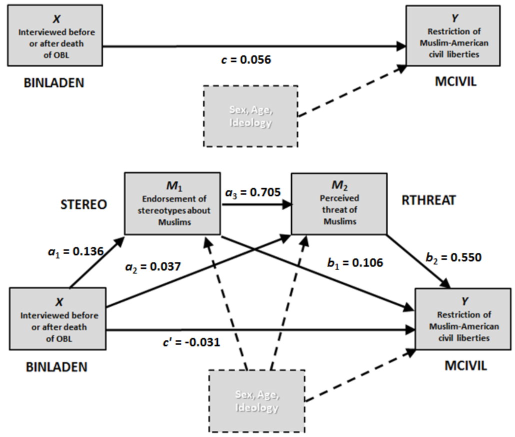


This model includes four pathways of influence of news coverage of OBL death, two through a single mediator, one through both mediators in serial, and one direct.

Example

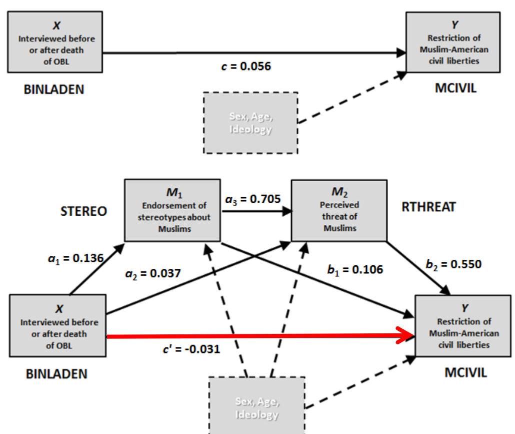


Example



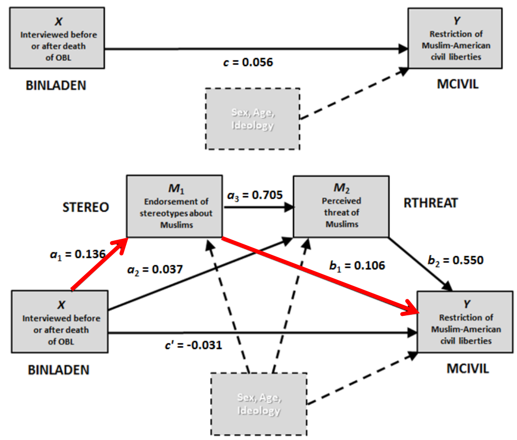
Direct effect = -0.031
Specific indirect effect via stereotype endorsement: 0.136(0.106) = 0.014
Specific indirect effect via perceived threat: 0.037(0.550) = 0.020
Specific indirect effect via stereotype endorsement and threat: 0.136(0.705)(0.550) = 0.053
Total indirect effect = 0.014 + 0.020 + 0.053 = 0.087
Total effect = -0.031 + 0.087 = 0.056

Example



Direct effect = -0.031
Specific indirect effect via stereotype endorsement: 0.136(0.106) = 0.014
Specific indirect effect via perceived threat: 0.037(0.550) = 0.020
Specific indirect effect via stereotype endorsement and threat: 0.136(0.705)(0.550) = 0.053
Total indirect effect = 0.014 + 0.020 + 0.053 = 0.087
Total effect = -0.031 + 0.087 = 0.056

Example



Direct effect = -0.031

Specific indirect effect via stereotype endorsement: 0.136(0.106) = 0.014

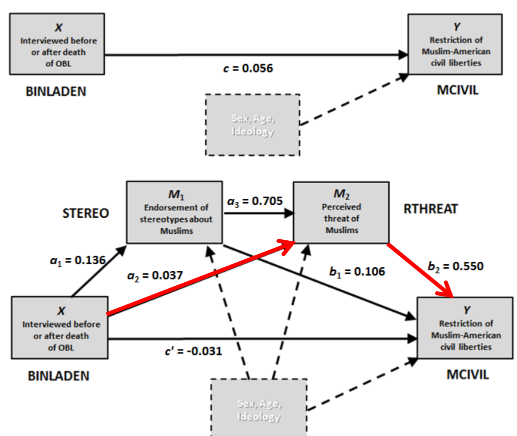
Specific indirect effect via perceived threat: 0.037(0.550) = 0.020

Specific indirect effect via stereotype endorsement and threat: 0.136(0.705)(0.550) = 0.053

Total indirect effect = 0.014 + 0.020 + 0.053 = 0.087

Total effect = -0.031 + 0.087 = 0.056

Example



Direct effect = -0.031

Specific indirect effect via stereotype endorsement: 0.136(0.106) = 0.014

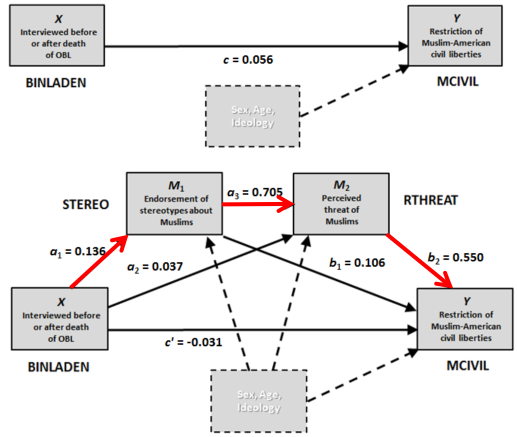
Specific indirect effect via perceived threat: 0.037(0.550) = 0.020

Specific indirect effect via stereotype endorsement and threat: 0.136(0.705)(0.550) = 0.053

Total indirect effect = 0.014 + 0.020 + 0.053 = 0.087

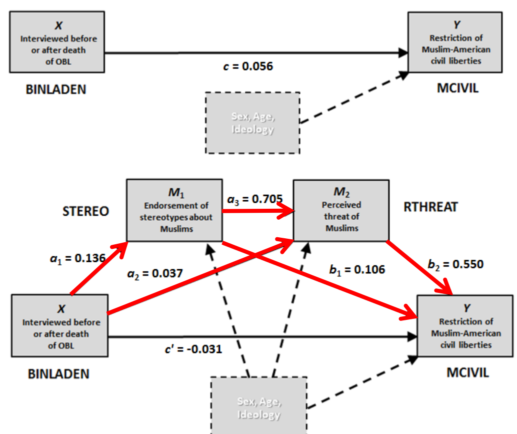
Total effect = -0.031 + 0.087 = 0.056

Example



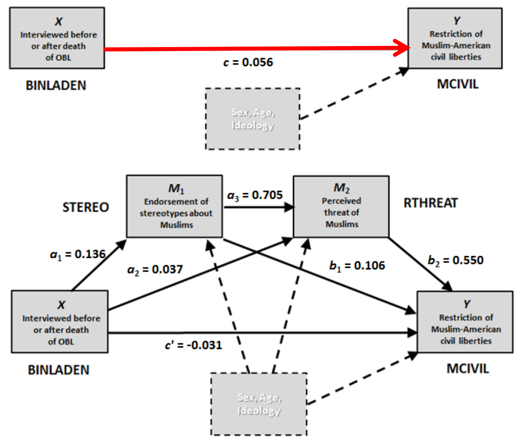
Direct effect = -0.031
Specific indirect effect via stereotype endorsement: 0.136(0.106) = 0.014
Specific indirect effect via perceived threat: 0.037(0.550) = 0.020
Specific indirect effect via stereotype endorsement and threat: 0.136(0.705)(0.550) = 0.053
Total indirect effect = 0.014 + 0.020 + 0.053 = 0.087
Total effect = -0.031 + 0.087 = 0.056

Example



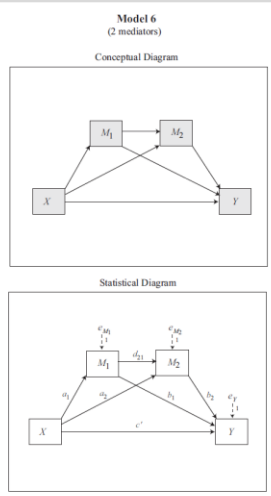
Direct effect = -0.031
Specific indirect effect via stereotype endorsement: 0.136(0.106) = 0.014
Specific indirect effect via perceived threat: 0.037(0.550) = 0.020
Specific indirect effect via stereotype endorsement and threat: 0.136(0.705)(0.550) = 0.053
Total indirect effect = 0.014 + 0.020 + 0.053 = 0.087
Total effect = -0.031 + 0.087 = 0.056

Example



Direct effect = -0.031
Specific indirect effect via stereotype endorsement: 0.136(0.106) = 0.014
Specific indirect effect via perceived threat: 0.037(0.550) = 0.020
Specific indirect effect via stereotype endorsement and threat: 0.136(0.705)(0.550) = 0.053
Total indirect effect = 0.014 + 0.020 + 0.053 = 0.087
Total effect = -0.031 + 0.087 = 0.056

Estimation and inference using PROCESS



PROCESS model 6 is the serial multiple mediator model.

In model 6, order of the variables in the "m=" list matters. Variables listed earlier are causally prior to those listed later. PROCESS allows up to four mediators to be linked in a causal chain. All possible indirect and direct effects are estimated.

```
process cov=sex age ideo/y=mcivil/x=binladen/m=stereo rthreat /boot=10000/model=6/total=1.
```

```
%process (data=binladen,cov=sex age ideo,y=mcivil,x=binladen,m=stereo rthreat,boot=10000,
          model=6,total=1);
```

PROCESS output

Output F

```
**************************************************************************
Model = 6
    Y = mcivil
    X = binladen
   M1 = stereo
   M2 = rthreat

Statistical Controls:
CONTROL= sex      age      ideo

Sample size
        661
**************************************************************************
Outcome: stereo

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .3557      .1265      .6495    23.7609     4.0000   656.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.9045      .1322    14.4084      .0000     1.6449     2.1640
binladen      .1358      .0639     2.1258      .0339      .0104      .2613
sex           .0398      .0635      .6262      .5314     -.0849      .1644
age           .0504      .0192     2.6220      .0089      .0127      .0882
ideo          .1293      .0143     9.0483      .0000      .1012      .1574

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

$$\hat{M}_1 = 1.905 + 0.136X + \ldots$$

a_1 path

PROCESS output

Output F

```
**************************************************************************
Outcome: rthreat

          R       R-sq        MSE          F        df1        df2          p
      .6764      .4575      .6076   110.4916     5.0000   655.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     -.2548      .1467    -1.7369      .0829     -.5428      .0332
stereo        .7047      .0378    18.6630      .0000      .6306      .7789
binladen      .0374      .0620      .6038      .5462     -.0843      .1592
sex           .1286      .0614     2.0938      .0367      .0080      .2492
age           .0451      .0187     2.4135      .0161      .0084      .0818
ideo          .0898      .0147     6.1257      .0000      .0610      .1186

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

$$\hat{M}_2 = -0.255 + 0.037X + 0.705M_1 + \ldots$$

a_3 path

a_2 path

PROCESS output

Output F

```
**************************************************************************
Outcome: mcivil

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .6727      .4526      .5890    90.1100     6.0000   654.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant      .7165      .1448     4.9499      .0000      .4323     1.0008
stereo        .1057      .0460     2.2965      .0220      .0153      .1960
rthreat       .5491      .0385    14.2732      .0000      .4736      .6247
binladen     -.0311      .0611     -.5095      .6106     -.1510      .0888
sex          -.1001      .0607    -1.6504      .0993     -.2193      .0190
age          -.0103      .0185     -.5599      .5758     -.0466      .0259
ideo          .0545      .0148     3.6696      .0003      .0253      .0836

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

$$\boxed{\hat{Y} = 0.717 - 0.031X + 0.106M_1 + 0.549M_2 + \ldots}$$

b_1 **path** (stereo)

b_2 **path** (rthreat)

c' **path** (binladen)

PROCESS output

Output F

```
************************** TOTAL EFFECT MODEL ****************************
Outcome: mcivil

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .3675      .1351      .9278    25.6100     4.0000   656.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.5149      .1580     9.5894      .0000     1.2047     1.8251
binladen      .0564      .0764      .7380      .4608     -.0936      .2063
sex          -.0099      .0759     -.1310      .8958     -.1589      .1391
age           .0393      .0230     1.7085      .0880     -.0059      .0844
ideo          .1675      .0171     9.8053      .0000      .1339      .2010

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

$$\boxed{\hat{Y} = 1.515 + 0.056X + \ldots}$$

c **path** (binladen)

PROCESS output

Output E

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

Total effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .0564      .0764      .7380      .4608     -.0936      .2063

Direct effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
     -.0311      .0611     -.5095      .6106     -.1510      .0888

Indirect effect(s) of X on Y:
           Effect    BootSE  BootLLCI  BootULCI
TOTAL       .0875     .0460     .0005     .1798
Ind1        .0144     .0098    -.0002     .0373
Ind2        .0206     .0338    -.0446     .0895
Ind3        .0526     .0248     .0050     .1021

Indirect effect key
 Ind1 :   binladen ->       stereo   ->       mcivil
 Ind2 :   binladen ->       stereo   ->       rthreat  ->       mcivil
 Ind3 :   binladen ->       rthreat  ->       mcivil
```

***c* path** (Total effect)

***c'* path** (Direct effect)

TOTAL: $a_1b_1 + a_2b_2 + a_1a_3b_2$

Ind1: a_1b_1

Ind2: $a_1a_3b_2$

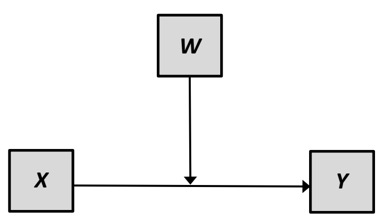
Ind3: a_2b_2

and bootstrap CIs

The data are consistent with the claim that coverage of OBL's death increased endorsement of restriction of Muslim civil liberties serially through stereotype endorsement and perceived threat of Muslims (0.053, 95% CI=0.005 to 0.102) but not through stereotype endorsement independent of perceived threat (.014, 95% CI = -0.0002 to 0.037) or perceived threat independent of stereotype endorsement (0.021, 95% CI = -0.045 to 0.089). There is no evidence of a direct effect of his death independent of these pathways of influence.

Moderation

Moderation. The effect of *X* on *Y* can be said to be *moderated* if its size or direction is dependent on some third variable *W*. It tells us about the conditions that facilitate, enhance, or inhibit the effect, or for whom or what the effect is large vs. small, present vs. absent, positive vs. negative vs. zero.



In this diagram, *W* is depicted to *moderate* the size of the effect of *X* on *Y*, meaning that the size of the effect of *X* on *Y* depends on *W*. In such a case, we say *W* is the *moderator* of the *X* → *Y* relationship, or that *X* and *W* *interact* in their influence on *Y*. *X* is sometimes called the focal predictor, and *W* the moderator.

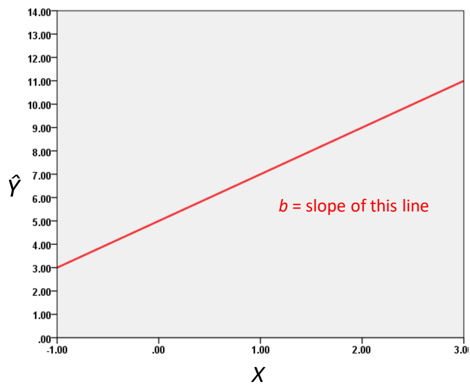
The simple regression coefficient

Consider a simple regression model with predictor variable *X*.

$$\hat{Y} = i_1 + bX \quad \text{such as} \quad \hat{Y} = 5.00 + 2.00X$$

Two cases that differ by one unit on *X* are estimated to differ by b = 2.00 units on *Y*. b is a "global property" of the model, in that makes no difference which value of *X* you start at---b is the estimated difference in *Y* between two cases who differ by a unit on *X*. Most generally, $b = \hat{Y}|(X = \omega + 1) - \hat{Y}|(X = \omega)$ for all ω.

X	$\hat{Y}$
-1	3.00
0	5.00
1	7.00
2	9.00
3	11.00



Partial regression coefficients as unconditional effects

Consider a multiple regression model with two predictors, *X* and *W*.

$$\hat{Y} = i + b_1X + b_2W \quad \text{such as} \quad \hat{Y} = 4.50 + 2.00X + 0.50W$$

Regardless of *W*, a one unit difference in *X* is associated with the same expected difference on *Y*. And regardless of the value of *X*, a one unit difference in *W* is associated with the same expected difference on *Y*. This is true regardless of which value of *X* or *W* you choose. b_1 and b_2 are global properties of the model.

Most generally,

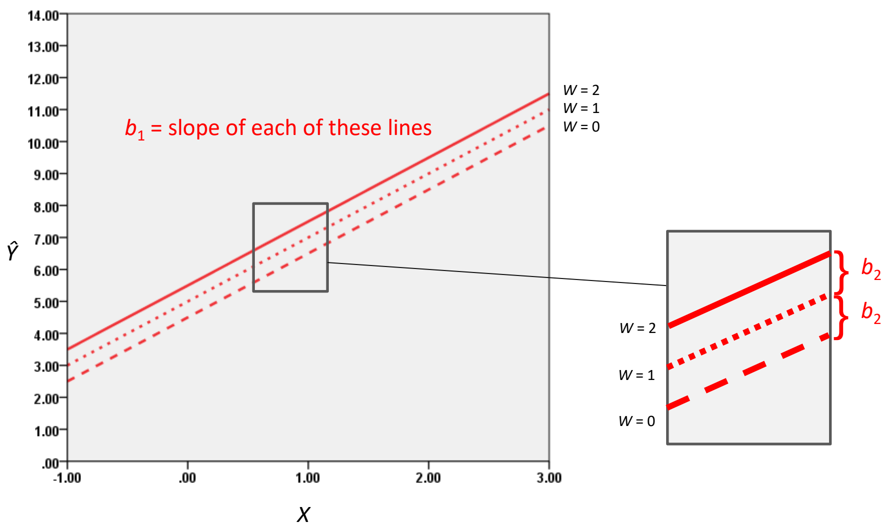
$b_1 = \hat{Y}|(X = \omega + 1, W = \lambda) - \hat{Y}|(X = \omega, W = \lambda)$ for all ω, λ

$b_2 = \hat{Y}|(W = \lambda + 1, X = \omega) - \hat{Y}|(W = \lambda, X = \omega)$ for all λ, ω

X	*W*	$\hat{Y}$
-1	0	2.50
-1	1	3.00
-1	2	3.50
0	0	4.50
0	1	5.00
0	2	5.50
1	0	6.50
1	1	7.00
1	2	7.50
2	0	8.50
2	1	9.00
2	2	9.50

Partial regression coefficients as unconditional effects

$$\hat{Y} = 4.50 + 2.00X + 0.50W$$



Releasing this constraint on the model

Suppose we let X's effect be a function of W, $f(W)$, as in

$$\hat{Y} = i + f(W)X + b_2W$$

For instance, let $f(W)$ be a linear function of W, $b_1 + b_3W$. Thus,

$$\hat{Y} = i + (b_1 + b_3W)X + b_2W$$

This can be rewritten in an equivalent form as

$$\hat{Y} = i + b_1X + b_2W + b_3XW$$

This model, the "simple moderation model," allows X's effect on Y to depend linearly on W. Other forms of moderation are possible, but this form is the one most frequently estimated.

Moderation

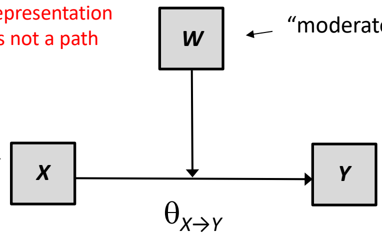
$$\hat{Y} = i + b_1X + b_2W + b_3XW$$

can be written as

$$\hat{Y} = i + (b_1 + b_3W)X + b_2W$$

This is a conceptual representation of moderation. This is not a path diagram.

“focal predictor”



“moderator”

$\theta_{X\to Y} = b_1 + b_3W$

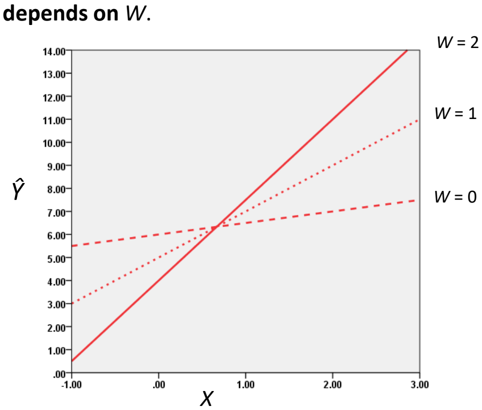
$\boxed{\theta_{X\to Y}\text{ is the “conditional effect of } X\text{” defined by the function } b_1 + b_3W}$

X’s effect as a function of *M*

i = 6.00
b_1 = 0.50
b_2 = -1.00
b_3 = 1.50

$$\hat{Y} = 6.00 + 0.50X - 1.00W + 1.50XW$$

Observe that the amount by which two cases that differ by one unit on *X* are estimated to differ on *Y* **depends on** *W*.



X	*W*	$\hat{Y}$
-1	0	5.50
-1	1	3.00
-1	2	0.50
0	0	6.00
0	1	5.00
0	2	4.00
1	0	6.50
1	1	7.00
1	2	7.50
2	0	7.00
2	1	9.00
2	2	11.00

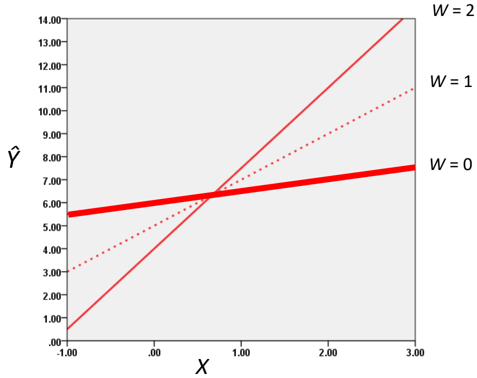
X's effect as a function of M

i = 6.00
b_1 = 0.50
b_2 = -1.00
b_3 = 1.50

$$\hat{Y} = 6.00 + 0.50X - 1.00W + 1.50XW$$

X	W	$\hat{Y}$
-1	0	5.50
-1	1	3.00
-1	2	0.50
0	0	6.00
0	1	5.00
0	2	4.00
1	0	6.50
1	1	7.00
1	2	7.50
2	0	7.00
2	1	9.00
2	2	11.00

Observe that the amount by which two cases that differ by one unit on X are estimated to differ on Y **depends on** W.



$$\theta_{X \to Y} = b_1 + b_3W$$
$$= 0.50 + 1.50W$$
$$= 0.50 + 1.50(0) = 0.50$$

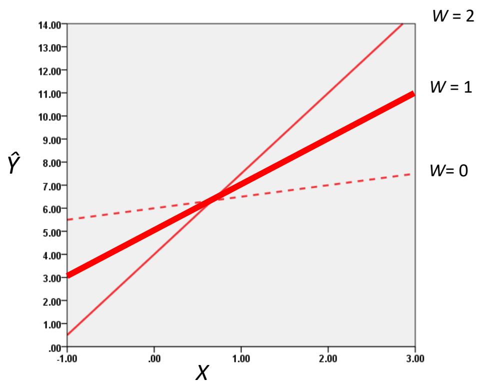
X's effect as a function of M

i = 6.00
b_1 = 0.50
b_2 = -1.00
b_3 = 1.50

$$\hat{Y} = 6.00 + 0.50X - 1.00W + 1.50XW$$

X	W	$\hat{Y}$
-1	0	5.50
-1	1	3.00
-1	2	0.50
0	0	6.00
0	1	5.00
0	2	4.00
1	0	6.50
1	1	7.00
1	2	7.50
2	0	7.00
2	1	9.00
2	2	11.00

Observe that the amount by which two cases that differ by one unit on X are estimated to differ on Y **depends on** W.



$$\theta_{X \to Y} = b_1 + b_3W$$
$$= 0.50 + 1.50W$$
$$= 0.50 + 1.50(1) = 2.00$$

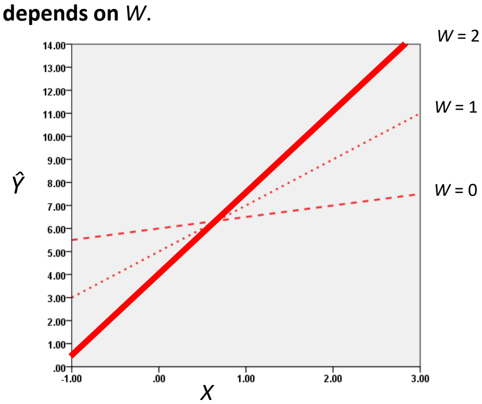
X's effect as a function of M

i = 6.00
b_1 = 0.50
b_2 = -1.00
b_3 = 1.50

$$\hat{Y} = 6.00 + 0.50X - 1.00W + 1.50XW$$

X	W	$\hat{Y}$
-1	0	5.50
-1	1	3.00
-1	2	0.50
0	0	6.00
0	1	5.00
0	2	4.00
1	0	6.50
1	1	7.00
1	2	7.50
2	0	7.00
2	1	9.00
2	2	11.00

Observe that the amount by which two cases that differ by one unit on X are estimated to differ on Y **depends on** W.



$$\theta_{X \to Y} = b_1 + b_3W$$
$$= 0.50 + 1.50W$$
$$= 0.50 + 1.50(2) = 3.50$$

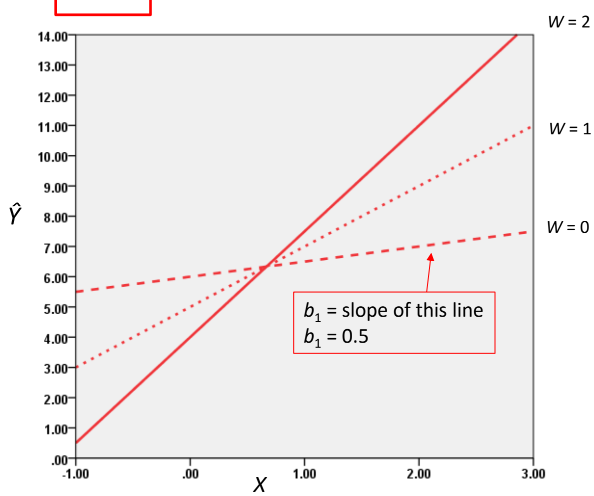
Interpretation of b_1 as a conditional effect

i = 6.00
b_1 = 0.50
b_2 = -1.00
b_3 = 1.50

$$\hat{Y} = 6.00 + \boxed{0.50X} - 1.00W + 1.50XW$$

b_1 is the effect of X on Y when W = 0. It quantifies how much two cases that differ by one unit on X but with W = 0 are estimated to differ on Y.

b_1 is a local property of the model. It characterizes the association between X and Y only when W = 0.



$$\boxed{b_1 = \hat{Y}|(X = \omega + 1, W = 0) - \hat{Y}|(X = \omega, W = 0) \text{ for all } \omega.}$$

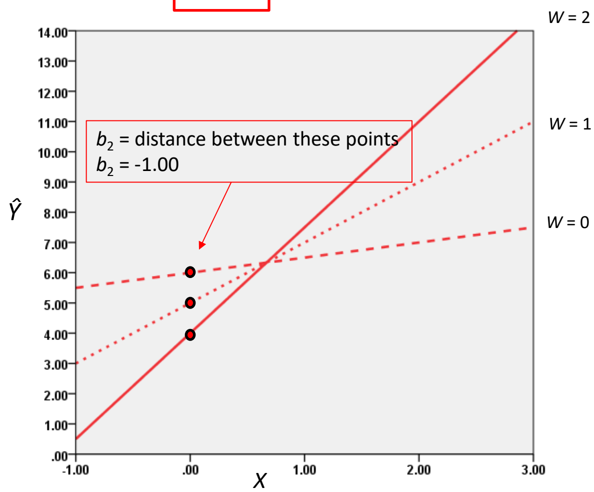
Interpretation of b_2 as a conditional effect

i = 6.00
b_1 = 0.50
b_2 = -1.00
b_3 = 1.50

$$\hat{Y} = 6.00 + 0.50X \boxed{-1.00}W + 1.50XW$$

b_2 is the effect of W when X = 0. It quantifies how much two cases that differ by one unit on W but with X = 0 are estimated to differ on Y.

b_2 is a local property of the model. It characterizes the association between W and Y only when X = 0.



$$\boxed{b_2 = \hat{Y}|(W = \lambda + 1, X = 0) - \hat{Y}|(W = \lambda, X = 0) \text{ for all } \lambda}$$

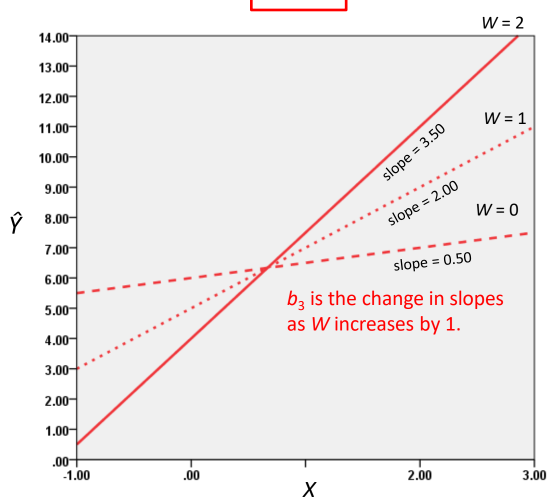
Interpretation of b_3

$$\hat{Y} = 6.00 + 0.50X - 1.00W \boxed{+1.50}XW$$

b_3 is the amount by which the conditional effect of X changes as M changes by one unit.

$\theta_{X\to Y} = b_1 + b_3W$
$= 0.50 + 1.50W$

$\theta_{X\to Y}$	W
0.50	0
2.00	1
3.50	2



$$b_3 = (\theta_{X\to Y}|W = \lambda + 1) - (\theta_{X\to Y}|W = \lambda) \text{ for all } \lambda$$

Differences in interpretation

	$\hat{Y} = i + b_1X + b_2W$	$\hat{Y} = i + b_1X + b_2W + b_3XW$
i	The estimated value of *Y* when *X* and *W* = 0.	The estimated value of *Y* when *X* and *W* = 0.
b_1	The effect of *X* on *Y* holding *W* constant. This is a *partial* effect.	The effect of *X* on *Y* when *W* = 0. This is a *conditional* effect. It is Not a “main effect” or “average effect” of *X*.
b_2	The effect of *W* on *Y* holding *X* constant. This is a *partial* effect.	The effect of *W* on *Y* when *X* = 0. This is a *conditional* effect. It is not a “main effect” or “average effect” of *W*.
b_3		How much the effect of *X* on *Y* changes as *W* changes by 1 unit.

Symmetry in moderation

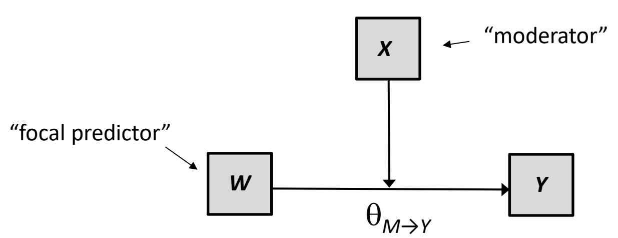
$$\boxed{\hat{Y} = i + b_1X + b_2W + b_3XW}$$

We saw that this is alternative representation of

$$\hat{Y} = i + (b_1 + b_3W)X + b_2W$$

But it is also an alternative representation of

$$\hat{Y} = i + (b_2 + b_3X)W + b_1X$$

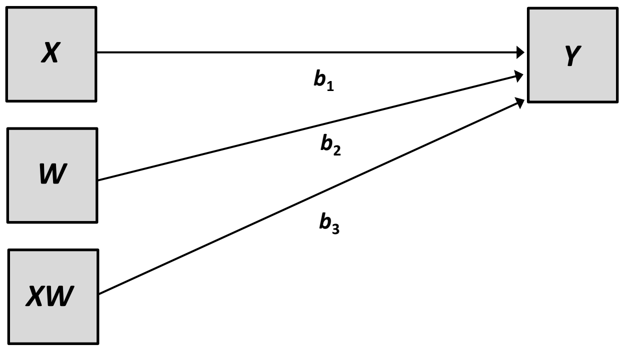


$\theta_{W \rightarrow Y} = b_2 + b_3X$

Here, *X* moderates the size of the effect of *W* on *Y*. Now *X* is the moderator. Ultimately, which variable *X* or *W* we think of as the moderator depends on substantive concerns. Statistically, it makes no difference as they are mathematically equivalent models.

In path diagram form

$$\hat{Y} = i + b_1X + b_2W + b_3XW$$

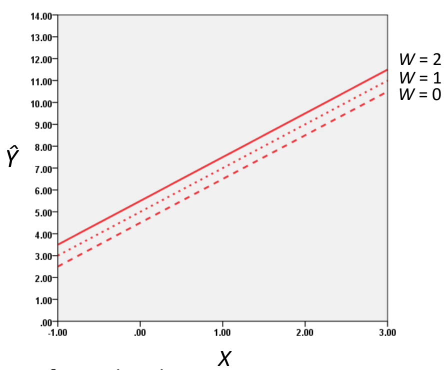


Remember

b_1 is NOT the effect of X on Y. The effect of X is $b_1 + b_3W$
b_2 is NOT the effect of W on Y. The effect of W is $b_2 + b_3X$

The importance of b_3 when testing a moderation hypothesis

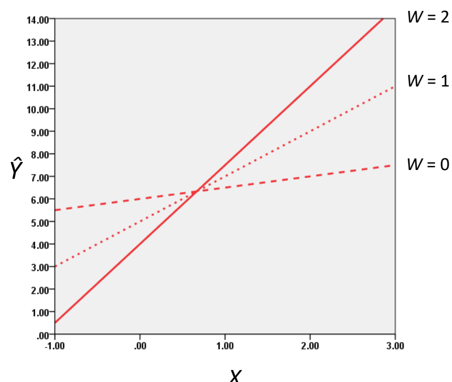
$$\hat{Y} = 4.50 + 2.00X + 0.50W + 0XW$$



$\theta_{X\to Y} = b_1 + b_3W$
$= 2.00 + 0W$

$\theta_{W\to Y} = b_2 + b_3X$
$= 0.50 + 0X$

$$\hat{Y} = 6.00 + 0.50X - 1.00W + 1.50XW$$



$\theta_{X\to Y} = b_1 + b_3W$
$= 0.50 + 1.50W$

$\theta_{W\to Y} = b_2 + b_3X$
$= -1.00 + 1.50X$

When $b_3 = 0$, a one unit change in X has the same effect on Y regardless of W, and a one unit change in W has the same effect on Y regardless of X. When $b_3 \neq 0$, the effect of a change in X on Y depends on W, and the effect of a change in W on Y depends on X. So we test a moderation hypothesis by testing whether b_3 is different from zero.

Example inspired by …

Witkiewitz, K., & Bowen, S. (2010). Depression, craving, and substance use following a randomized trial of mindfulness-based relapse prevention. ***Journal of Consulting and Clinical Psychology, 78***, 362-374.

Depression, Craving, and Substance Use Following a Randomized Trial of Mindfulness-Based Relapse Prevention

Katie Witkiewitz
Washington State University Vancouver

Sarah Bowen
University of Washington

168 clients of a public service agency providing treatment for alcohol and substance use disorders.

MBRP : Randomly assigned to treatment as usual (0) or mindfulness-based relapse prevention therapy (1)

BDI0: Beck Depression Inventory at start of therapy (0 to 3; multiply by 21 to see BDI in its original 0 to 63 metric). This is also available at the termination of therapy (**BDIP**)

CRAVE2: Score on the Penn Alcohol Craving Scale at 2 month follow-up (0 to 6). Also available at baseline, prior to start of therapy (**CRAVE0**)

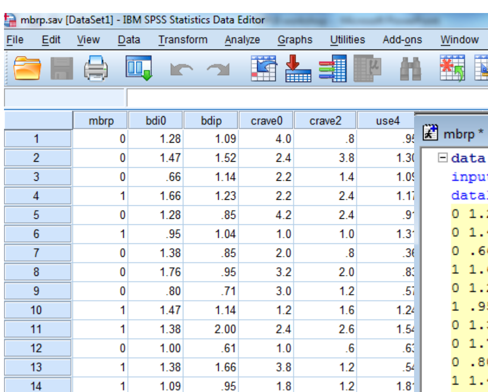
USE4: Alcohol and other substance use at 4-month follow-up. (0 to 5)

TREATHRS: Hours of therapy administered.

The data file is MBRP

The Data: MBRP

SPSS



	mbrp	bdi0	bdip	crave0	crave2	use4
1	0	1.28	1.09	4.0	.8	.9
2	0	1.47	1.52	2.4	3.8	1.3
3	0	.66	1.14	2.2	1.4	1.0
4	1	1.66	1.23	2.2	2.4	1.1
5	0	1.28	.85	4.2	2.4	.9
6	1	.95	1.04	1.0	1.0	1.3
7	0	1.38	.85	2.0	.8	.3
8	0	1.76	.95	3.2	2.0	.8
9	0	.80	.71	3.0	1.2	.5
10	1	1.47	1.14	1.2	1.6	1.2
11	1	1.38	2.00	2.4	2.6	1.5
12	0	1.00	.61	1.0	.6	.6
13	1	1.38	1.66	3.8	1.2	.5
14	1	1.09	.95	1.8	1.2	1.8

SAS

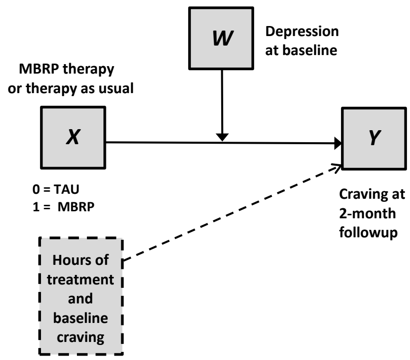
mbrp *

```
data mbrp;
input mbrp bdi0 bdip crave0 crave2 use4 treathrs;
datalines;
0 1.28 1.09 4.0 .8 .95 32
0 1.47 1.52 2.4 3.8 1.30 36
0 .66 1.14 2.2 1.4 1.09 34
1 1.66 1.23 2.2 2.4 1.17 37
0 1.28 .85 4.2 2.4 .91 27
1 .95 1.04 1.0 1.0 1.31 32
0 1.38 .85 2.0 .8 .36 38
0 1.76 .95 3.2 2.0 .83 26
0 .80 .71 3.0 1.2 .57 42
1 1.47 1.14 1.2 1.6 1.24 32
1 1.38 2.00 2.4 2.6 1.54 25
0 1.00 .61 1.0 .6 .63 24
1 1.38 1.66 3.8 1.2 .54 35
1 1.09 .95 1.8 1.2 1.81 34
```

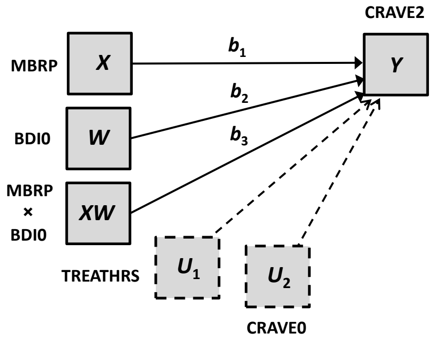
These aren't their actual data. But the analyses we do yield similar results to what they report.

Example

Conceptual diagram



Statistical model



Does the effect of MBRP therapy relative to therapy as usual on craving depend on initial depression? That is, is the therapy more or less effective as a function of depression prior to start of therapy?

Estimation using OLS regression

```
compute mbrpdep = mbrp*bdi0.
regression/dep = crave2/method = enter mbrp bdi0 mbrpdep treathrs crave0.
```

```
data mbrp;set mbrp;mbrpdep=mbrp*bdi0;run;
proc reg data=mbrp;model crave2=mbrp bdi0 mbrpdep treathrs crave0;run;
```

***X* = MBRP**
***W* = BDI0**
***Y* = CRAVE2**

$$\hat{Y} = 1.038 + 0.587X + 1.122W - 0.948XW + \ldots$$

Output G

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.038	.470		2.209	.029
	MBRP: Therapy as usual (0) or MBRP therapy (1)	.587	.524	.299	1.120	.264
	BDI0: Beck Depression Inventory baseline	1.122	.276	.366	4.063	.000
	mbrpdep	-.948	.423	-.598	-2.240	.026
	TREATHRS: Hours of therapy	-.018	.010	-.120	-1.719	.088
	CRAVE0: Baseline craving	.192	.073	.183	2.614	.010

a. Dependent Variable: CRAVE2: Craving at two month follow-up

"conditional effects", not "main effects"

b_1 = 0.587
b_2 = 1.122
b_3 = -0.948

The coefficient for the product is statistically different from zero. This means that the effect of MBRP therapy on craving depends on the person's level of depression at the start of therapy. But to really understand what is happening, we need a picture.

Visualizing the model

Rejecting the null hypothesis that "true b_3" is equal to zero tells you that the focal predictor's effect is indeed moderated by the proposed moderator. But moderation can take many different forms. We need to visualize the effect in order to interpret the result.

Step 1: Select various combinations of values of the focal predictor and moderator. The selection is sometimes arbitrary, but it may not be. Just make sure the values chosen are within the range of the data.

Step 2: Using the model, generate the estimates of Y using your selected values of the focal predictor and moderator. If your model includes covariates, use the sample mean for each of those.

Step 3: Graph, using whatever graphics program you prefer.

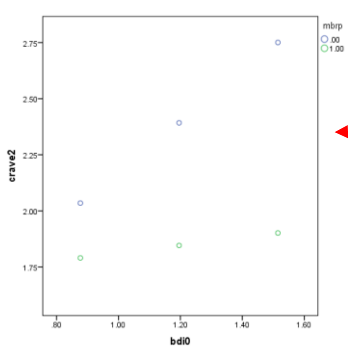
$$\hat{Y} = 1.038 + 0.587X + 1.122W - 0.948XW - 0.018U_1 + 0.192U_2$$

MBRP (X)	**BDI0** (W)	**TREATHRS** (U_1)	**CRAVE0**(U_2)	$\hat{Y}$
0	0.877	30.685	2.943	2.035
0	1.196	30.685	2.943	2.393
0	1.515	30.685	2.943	2.751
1	0.877	30.685	2.943	1.790
1	1.196	30.685	2.943	1.846
1	1.515	30.685	2.943	1.901

I used one standard deviation below the mean, the mean, and one standard deviation above the mean. It really makes no difference what you choose, except you want to make sure that your resulting graph is not extrapolating beyond the available data.

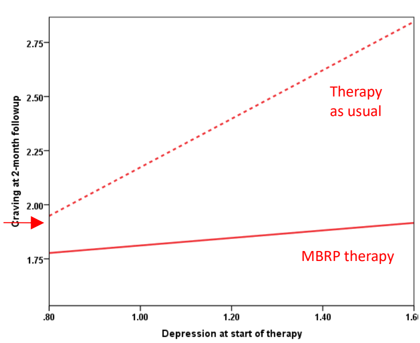
Example code in SPSS

```
data list free/mbrp bdi0.
begin data.
 0  0.877
 0  1.196
 0  1.515
 1  0.877
 1  1.196
 1  1.515
end data.
compute crave2=1.038+0.587*mbrp+1.122*bdi0-0.948*mbrp*bdi0-0.018*30.685+0.192*2.943.
graph/scatterplot = bdi0 with crave2 by mbrp.
```



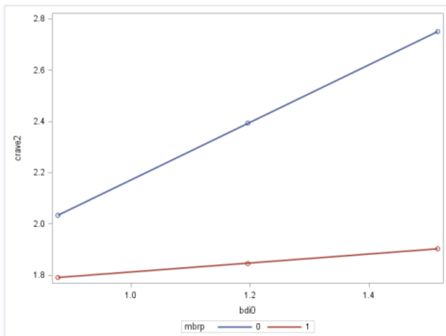
Before editing

After editing



Example code in SAS

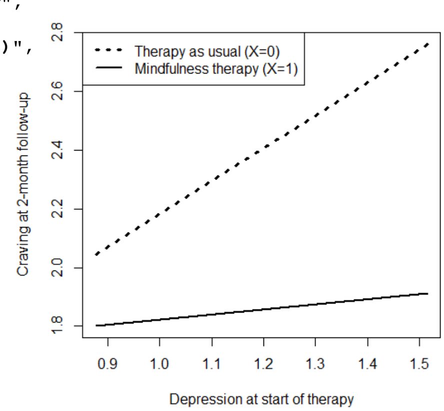
```
data;
input mbrp bdi0;
crave2=1.038+0.587*mbrp+1.122*bdi0-0.948*mbrp*bdi0-0.018*30.685+0.192*2.943;
datalines;
0  0.877
0  1.196
0  1.515
1  0.877
1  1.196
1  1.515
run;
proc sgplot;reg x=bdi0 y=crave2/group=mbrp;run;
```



Example code in R

Although hard to learn at first, once you learn how to use R, you will find it very helpful In the construction of visual depictions of models.

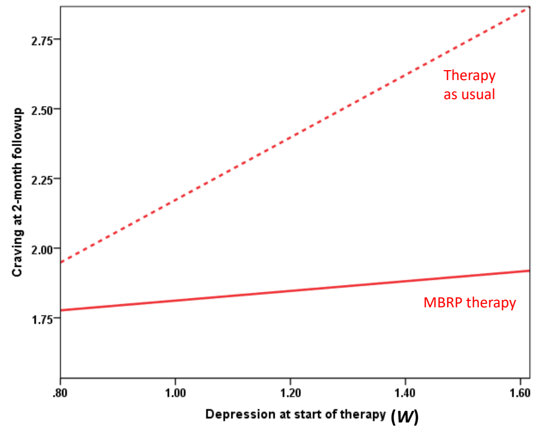
```
x<-c(0,1,0,1,0,1)
m<-c(0.877,0.877,1.196,1.196,1.515,1.515)
y<-1.038+0.587*x+1.122*m-0.948*x*m-0.018*30.685+0.192*2.943
plot(y=y,x=m,pch=15,col="white",
xlab="Depression at start of therapy",
ylab="Craving at 2-month follow-up")
legend.txt<-c("Therapy as usual (X=0)",
"Mindfulness therapy (X=1)")
legend("topleft",legend=legend.txt,
lty=c(3,1),lwd=c(3,2))
lines(m[x==0],y[x==0],lwd=3,lty=3)
lines(m[x==1],y[x==1],lwd=2,lty=1)
```



Substantive interpretation of the pattern

Therapy as usual (**MBRP** (X) =0)

MBRP therapy (**MBRP** (X) = 1)



Those who receive MBRP therapy crave substances less than those who receive MBRP therapy, but this difference is larger among those more depressed at the start of therapy.

A graphical depiction of the model

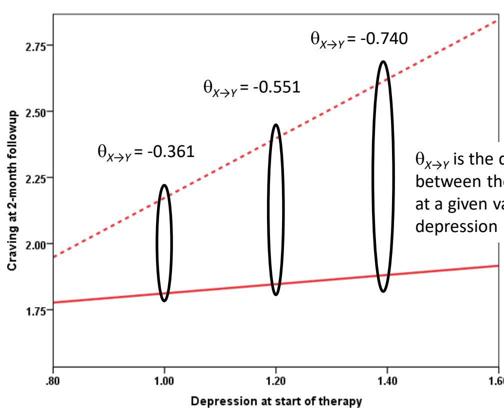
$$\hat{Y} = 1.038 + 0.587X + 1.122W - 0.948XW + \ldots$$

or, equivalently,

$$\hat{Y} = 1.038 + (0.587 - 0.948W)X + 1.122W + \ldots$$

Therapy as usual (**MBRP** (X) =0)

MBRP therapy (**MBRP** (X) = 1)



$\theta_{X\rightarrow Y}$ is the distance between the lines at a given value of depression

The conditional effect of MBRP therapy ($\theta_{X\rightarrow Y}$) is defined by the function

$$\theta_{X\rightarrow Y} = 0.587 - 0.948W$$

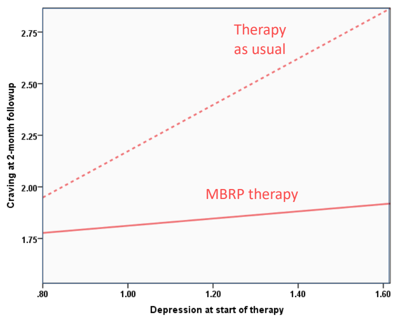
BDI0 (W)	$\theta_{X\rightarrow Y}$
1.00	-0.361
1.20	-0.551
1.40	-0.740

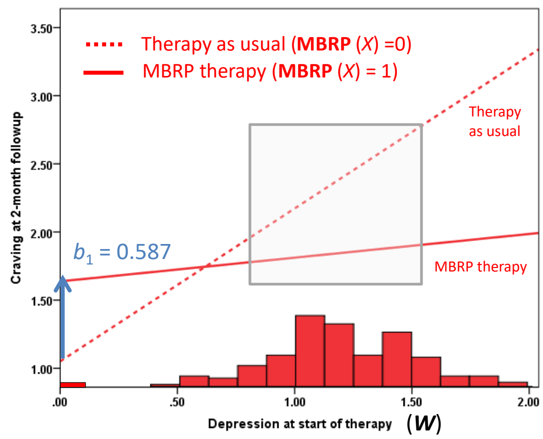
You can plug any value of BDI0 you want into the function to get the conditional effect of MBRP therapy

Interpretation of b_1

b_1 = 0.587

$$\hat{Y} = 1.038 + 0.587X + 1.122W - 0.948XW + \ldots$$





b_1 is the effect of X on Y when W = 0. It is a conditional effect, and a local term of the model.

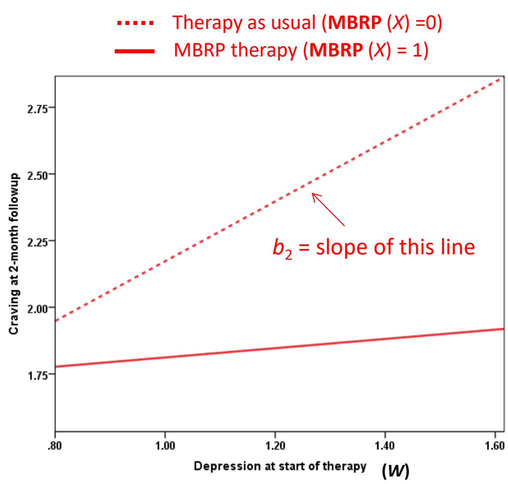
Interpretation of b_2

b_2 = 1.122

$$\hat{Y} = 1.038 + 0.587X + 1.122W - 0.948XW + \ldots$$

b_2 is the conditional effect of W when X = 0. It is a conditional effect and a local term of the model.

Among those given therapy as usual, those who were relatively more depressed at the start of therapy had relatively higher craving at two months follow-up



Interpretation of b_3

b_3 = -0.948

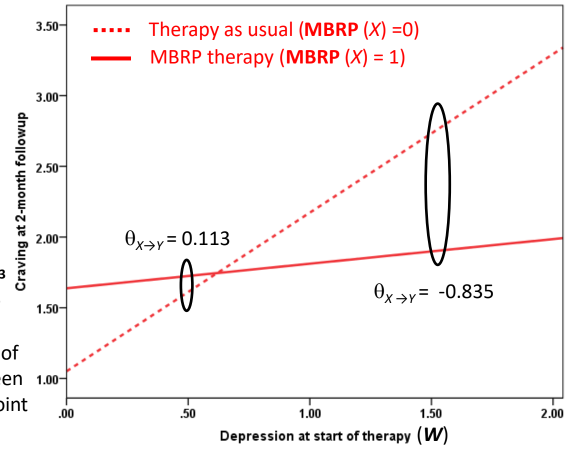
$$\hat{Y} = 1.038 + 0.587X + 1.122W - 0.948XW + \ldots$$

$\theta_{X\rightarrow Y}$ = 0.587 – 0.948W

W	$\theta_{X\rightarrow Y}$
0.50	**0.113**
1.00	-0.361
1.50	**-0.835**
2.00	-1.309

-0.835 – (0.113) = -0.948 = b_3
-1.309 – (-0.361) = -0.948 = b_3

b_3 is the difference in the effect of MBRP therapy on craving between those who differ by one scale point in their pre-therapy depression.



$\theta_{X\rightarrow Y} | (W = \lambda+1) - \theta_{X\rightarrow Y} | (W = \lambda)$ = -0.948, for all λ

Interpretation of b_3

b_3 = -0.948

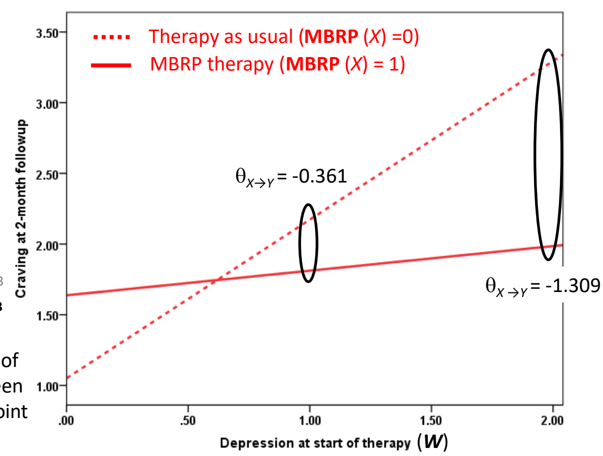
$$\hat{Y} = 1.038 + 0.587X + 1.122W - 0.948XW + \ldots$$

$\theta_{X\rightarrow Y}$ = 0.587 – 0.948W

W	$\theta_{X\rightarrow Y}$
0.50	0.113
1.00	**-0.361**
1.50	-0.835
2.00	**-1.309**

-0.835 – (0.113) = -0.948 = b_3
-1.309 – (-0.361) = -0.948 = b_3

b_3 is the difference in the effect of MBRP therapy on craving between those who differ by one scale point in their pre-therapy depression.



$\theta_{X\rightarrow Y} | (W=\lambda+1) - \theta_{X\rightarrow Y} | (W = \lambda)$ = -0.948, for all λ

Probing an interaction

The coefficient for the product term carries information about how changes in one variable are related to changes in the effect of the other. A picture helps to understand how the focal variable's effect changes as a function of the moderator variable.

It is typically desirable to conduct statistical tests of the focal predictor variable's effect at values of the moderator. This allows you to make more definitive claims about where the focal predictor variables effect is zero versus where it is not.

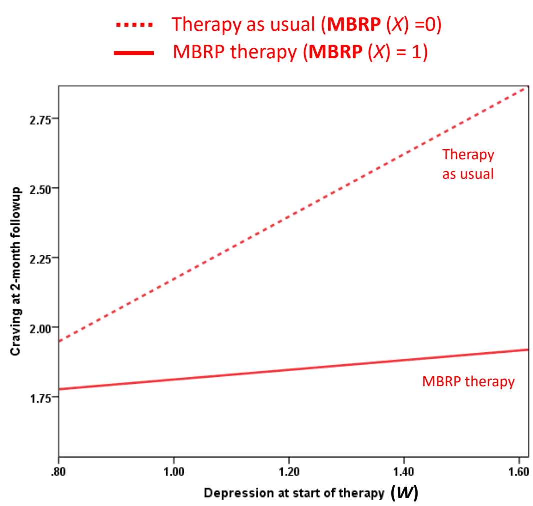
"Pick-a-Point" Approach

Select values of the moderator and estimate the conditional effect of the focal predictor at those values of the moderator, along with a hypothesis test or confidence interval.

Johnson-Neyman Technique

Derive mathematically where on the moderator variable continuum the focal variable's effect transitions between statistically significant and nonsignificant.

Substantive interpretation of the pattern



Those who receive MBRP therapy crave substances less than those who receive MBRP therapy, but this difference is larger among those more depressed at the start of therapy.

Pick-a-point approach

$$\hat{Y} = i + b_1X + b_2W + b_3XW$$

Select a value of the moderator (W) at which you'd like to have an estimate of $\theta_{X \rightarrow Y}$, the focal predictor variable's (X) effect. Then derive its standard error. The ratio of the effect to its standard error is distributed as $t(df_{residual})$ under the null hypothesis that the effect of the focal predictor is zero at that moderator value, where df_{residual} is the residual degrees of freedom from the regression model.

We already know that $\theta_{X \rightarrow Y} = b_1 + b_3W$

The estimated standard error of $\theta_{X \rightarrow Y}$ is

$$s_{\theta_{X \rightarrow Y}} = \sqrt{s_{b_1}^2 + 2Ws_{b_1b_3} + W^2 s_{b_3}^2}$$

Squared standard error of b_1

Covariance of b_1 and b_3

Squared standard error of b_3

You could do this by hand, and instructions are available in various books on regression analysis (e.g., Aiken and West, 1991; Cohen et al., 2003). But there is no reason to, and the potential for mistakes is high. It is made easier using "**regression centering**."

Pick-a-point: Regression centering approach

$$\hat{Y} = i + b_1X + b_2W + b_3XW$$

In the above model, b_1 estimates the conditional effect of X when $M = 0$. If we desire the conditional effect of X when M equals some value λ, we can produce a new variable M' that is M centered around λ, such that $M' = 0$ when $M = \lambda$. Then substitute M' for M in the model above. That is, we will estimate

$$\hat{Y} = i + b_1X + b_2(W - \lambda) + b_3X(W - \lambda)$$

as

$$\hat{Y} = i + b_1X + b_2W' + b_3XW' \quad \text{where} \quad W' = W - \lambda$$

In this model, b_1 is the conditional effect of X when $W' = 0$. But $W' = 0$ when $W = \lambda$. So b_1 estimates the conditional effect of X when $W = \lambda$. A common (but arbitrary) convention is to use $\lambda = \overline{W}$, $\lambda = \overline{W} - SD_W$, and $\lambda = \overline{W} + SD_W$

Pick-a-point: Regression centering approach

```
compute bdi0_p = bdi0-1.196.
compute interact = bdi0_p*mbrp.
regression/dep = crave2/method = enter mbrp bdi0_p interact treathrs crave0.
```

λ= 1.196
(the sample mean)

```
data mbrp;set mbrp;
bdi0_p=bdi0-1.196;
interact=bdi0_p*mbrp;
proc reg data=mbrp;model crave2=mbrp bdi0_p interact treathrs crave0;run;
```

Coefficients[a]

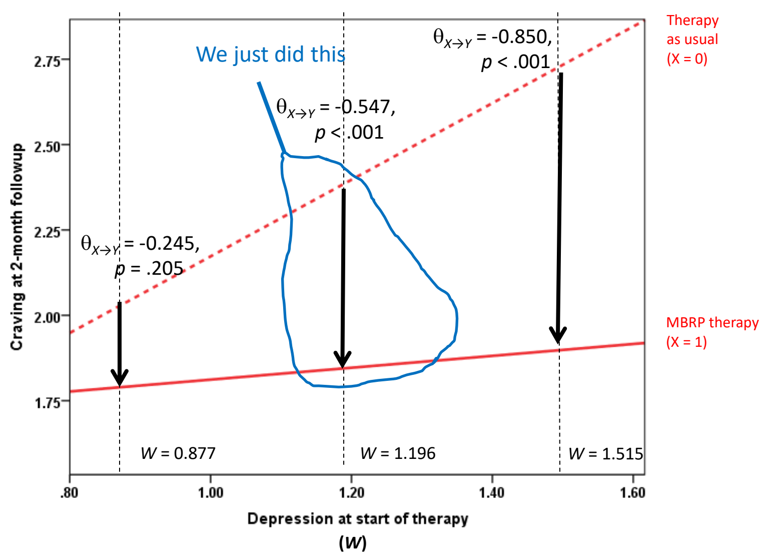
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.380	.364		6.534	.000
	MBRP: Therapy as usual (0) or MBRP therapy (1)	-.547	.137	-.279	$\boxed{-3.980}$	$\boxed{.000}$
	bdi0_p	1.122	.276	.366	4.063	.000
	interact	-.948	.423	-.197	-2.240	.026
	TREATHRS: Hours of therapy	-.018	.010	-.120	-1.719	.088
	CRAVE0: Baseline craving	.192	.073	.183	2.614	.010

a. Dependent Variable: CRAVE2: Craving at two month follow-up

$\theta_{X \to Y} | (W = 1.196)$ $\qquad s_{\theta_{X \to Y}}$

MBRP therapy reduces craving relative to therapy as usual among people "average" in pre-therapy depression, $\theta_{X \to Y}$ = -0.547, p < .001.

Repeat for other values of the moderator



Pick-a-point: Regression centering approach

```
compute bdi0_p = bdi0-0.877.
compute interact = bdi0_p*mbrp.
regression/dep = crave2/method = enter mbrp bdi0_p interact treathrs crave0.
```

λ= 0.877
(One SD below the sample mean)

```
data mbrp;set mbrp;
bdi0_p=bdi0-0.877;
interact=bdi0_p*mbrp;
proc reg data=mbrp;model crave2=mbrp bdi0_p interact treathrs crave0;run;
```

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.023	.367		5.506	.000
	MBRP: Therapy as usual (0) or MBRP therapy (1)	-.245	.192	-.124	-1.272	.205
	bdi0_p	1.122	.276	.366	4.063	.000
	interact	-.948	.423	-.242	-2.240	.026
	TREATHRS: Hours of therapy	-.018	.010	-.120	-1.719	.088
	CRAVE0: Baseline craving	.192	.073	.183	2.614	.010

a. Dependent Variable: CRAVE2: Craving at two month follow-up

$\theta_{X \to Y} | (W = 0.877)$ $\quad s_{\theta_{X \to Y}}$

MBRP therapy does not reduce craving relative to therapy as usual among people "relatively low" in pre-therapy depression, $\theta_{X \to Y} = -0.245$, $p = .21$.

Pick-a-point: Regression centering approach

```
compute bdi0_p = bdi0-1.515.
compute interact = bdi0_p*mbrp.
regression/dep = crave2/method = enter mbrp bdi0_p interact treathrs crave0.
```

λ= 1.515
(One SD above the sample mean)

```
data mbrp;set mbrp;
bdi0_p=bdi0-1.515;
interact=bdi0_p*mbrp;
proc reg data=mbrp;model crave2=mbrp bdi0_p interact treathrs crave0;run;
```

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.738	.382		7.165	.000
	MBRP: Therapy as usual (0) or MBRP therapy (1)	-.850	.193	-.433	-4.398	.000
	bdi0_p	1.122	.276	.366	4.063	.000
	interact	-.948	.423	-.257	-2.240	.026
	TREATHRS: Hours of therapy	-.018	.010	-.120	-1.719	.088
	CRAVE0: Baseline craving	.192	.073	.183	2.614	.010

a. Dependent Variable: CRAVE2: Craving at two month follow-up

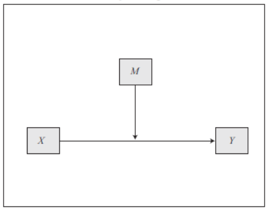
$\theta_{X \to Y} | (W = 1.515)$ $\quad s_{\theta_{X \to Y}}$

MBRP therapy reduces craving relative to therapy as usual among people "relatively high" in pre-therapy depression, $\theta_{X \to Y} = -0.850$, $p < .001$.

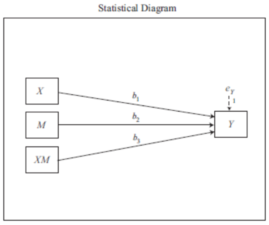
Using PROCESS

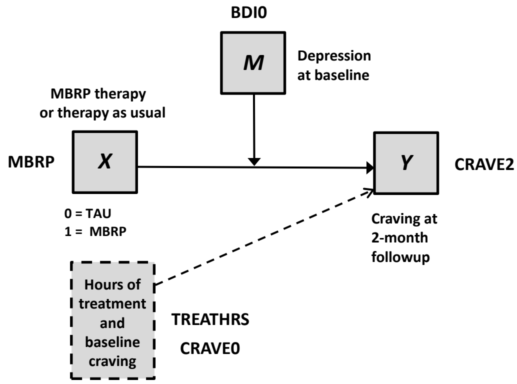
Model 1

Conceptual Diagram



Statistical Diagram






```
process cov = treathrs crave0/y=crave2/x=mbrp/w=bdi0
  /model=1/jn=1/plot=1/moments=1.
```

```
%process (data=mbrp,cov=treathrs crave0,y=crave2,x=mbrp,
   w=bdi0,model=1,jn=1,plot=1,moments=1);
```

PROCESS output

Output H

$$\hat{Y} = 1.038 + 0.587X + 1.122M - 0.948XM + \ldots$$

PROCESS generates the product term for you.

```
Model  : 1
    Y  : crave2
    X  : mbrp
    W  : bdi0

Covariates:
 treathrs crave0

Sample
Size:  168

**************************************************************************
OUTCOME VARIABLE:
 crave2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5140      .2642      .7277    11.6319     5.0000   162.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.0385      .4701     2.2090      .0286      .1102     1.9668
mbrp          .5872      .5241     1.1204      .2642     -.4478     1.6222
bdi0         1.1221      .2762     4.0625      .0001      .5767     1.6675
Int_1        -.9485      .4235    -2.2398      .0265    -1.7847     -.1122
treathrs     -.0177      .0103    -1.7190      .0875     -.0380      .0026
crave0        .1920      .0735     2.6138      .0098      .0470      .3371

Product terms key:
 Int_1    :        mbrp     x        bdi0

Test(s) of highest order unconditional interaction(s):
      R2-chng          F        df1        df2          p
X*W      .0228     5.0166     1.0000   162.0000      .0265
```

PROCESS output

PROCESS sees that the moderator is quantitative (because it has more than 2 values) so it automatically implements the pick-a-point procedure. When `moments = 1` moderator values equal to the mean of the moderator as well as ± one SD from the mean.

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

Conditional effect of X on Y at values of the moderator(s):
       bdi0     Effect         se          t          p       LLCI       ULCI
      .8772     -.2447      .1922    -1.2733      .2047     -.6243      .1348
     1.1963     -.5473      .1375    -3.9818      .0001     -.8188     -.2759
     1.5153     -.8500      .1933    -4.3973      .0000    -1.2317     -.4683

Values for quantitative moderators are the mean and plus/minus one SD from mean.
Values for dichotomous moderators are the two values of the moderator.

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

$\theta_{X\rightarrow Y} = 0.587 - 0.948W$

Output H

MBRP therapy resulted in lower craving than did therapy as usual among those relatively "moderate" ($\theta_{X\rightarrow Y|W=1.196}$ = -0.547, p < .001) or "relatively high" ($\theta_{X\rightarrow Y|W=1.515}$ = -0.850, p < .001) in pre-therapy depression. Among those "relatively low" in pre-therapy depression, MBRP therapy had no statistically significant effect on craving relative to therapy as usual. ($\theta_{X\rightarrow Y|W=0.877}$ = -0.245, p = .205)

PROCESS output: PLOT option

```
process cov = treathrs crave0/y=crave2/x=mbrp/w=bdi0
  /model=1/jn=1/plot=1/moments=1.
```

```
%process (data=mbrp,cov=treathrs crave0,y=crave2,x=mbrp,
   w=bdi0,model=1,jn=1,plot=1,moments=1);
```

Both the SPSS and SAS versions produce a table of estimated values of *Y* for different combinations of *X* and *M*. Plug these into your preferred graphing program to generate a plot, or use SPSS or SAS's graphics features. SPSS writes the code for you. Just cut and paste this into an SPSS syntax file and execute:

```
DATA LIST FREE/mbrp bdi0 crave2.
BEGIN DATA.
      .0000      .8772     2.0456
     1.0000      .8772     1.8009
      .0000     1.1963     2.4037
     1.0000     1.1963     1.8563
      .0000     1.5153     2.7617
     1.0000     1.5153     1.9117
END DATA.
GRAPH/SCATTERPLOT=bdi0 WITH crave2 BY mbrp.
```

Output H

Generating a graph from PROCESS "PLOT" option: SAS

```
data;
input mbrp bdi0 crave2;
datalines;
 .0000      .8772     2.0456
1.0000      .8772     1.8009
 .0000     1.1963     2.4037
1.0000     1.1963     1.8563
 .0000     1.5153     2.7617
1.0000     1.5153     1.9117
run;
proc sgplot;reg x=bdi0 y=crave2/group=mbrp;run;
```

Output generated by the PLOT option.

Generating a graph from PROCESS "PLOT" option: R

```
x<-c(0,1,0,1,0,1)
w<-c(0.877,0.877,1.196,1.196,1.515,1.515)
y<-c(2.046,1.801,2.404,1.856,2.762,1.912)
plot(y=y,x=w,pch=15,col="white",
xlab="Depression at start of therapy",
ylab="Craving at 2-month follow-up")
legend.txt<-c("Therapy as usual (X=0)","Mindfulness therapy (X=1)")
legend("topleft",legend=legend.txt,
lty=c(3,1),lwd=c(3,2))
lines(w[x==0],y[x==0],lwd=3,lty=3)
lines(w[x==1],y[x==1],lwd=2,lty=1)
```

From the PLOT option in PROCESS.

Additional probing options

Setting `moments = 0` or leaving it out, produces estimates of the conditional effect of *X* at the 16th, 50th, and 84th percentiles of the moderator rather than the mean and plus/minus one standard deviation. Or use the **mmodval** option to request a specific value of the moderator at which you'd like the conditional effect of *X.*

```
process vars = ... /moments = 0.
```

```
%process (data = ... , moments = 0);
```

Conditional effects of the focal predictor at values of the moderator(s):

bdi0	Effect	se	t	p	LLCI	ULCI
.9020	-.2683	.1850	-1.4500	.1490	-.6336	.0971
1.1900	-.5414	.1375	-3.9384	.0001	-.8129	-.2699
1.5180	-.8525	.1941	-4.3923	.0000	-1.2358	-.4692

W values in conditional tables are the 16th, 50th, and 84th percentiles.

```
process vars = ... /mmodval = 1.5.
```

```
%process (data = ... , mmodval = 1.5);
```

Conditional effect of X on Y at values of the moderators(s)

bdi0	Effect	se	t	p	LLCI	ULCI
1.5000	-.8354	.1888	-4.4253	.0000	-1.2082	-.4626

The Johnson-Neyman technique

The Johnson-Neyman technique seeks to find the value or values of the moderator *(W)* within the data, if they exist, such that the *p*-value for the ratio of the conditional effect of the focal predictor at that value or values of *W* is exactly equal to some chosen level of significance α

To do so, we ask what value of *W* produces a ratio exactly equal to the critical *t* value (t_{crit}) required to reject the null hypothesis that the conditional effect of *X* is equal to zero?

$$t_{crit} = \frac{b_1 + b_3 W}{\sqrt{\left\{s_{b_1}^2 + 2M s_{b_1 b_3}^2 + M^2 s_{b_3}^2\right\}}}$$

Isolate *W* and solve the polynomial that results. The quadratic formula finds the solutions:

$$W = \frac{-2(t_{crit}^2 s^{-}_{b_1 b_3} - b_1 b_3) \pm \sqrt{(2t_{crit}^2 s_{b_1 b_3} - 2b_1 b_3)^2 - 4(t_{crit}^2 s_{b_3}^2 - b_3^2)(t_{crit}^2 s_{b_1}^2 - b_1^2)}}{2(t_{crit}^2 s_{b_3}^2 - b_3^2)}$$

The Johnson-Neyman technique

This will produce no values, one value, or two values of *W* that are within the range of the moderator variable data.

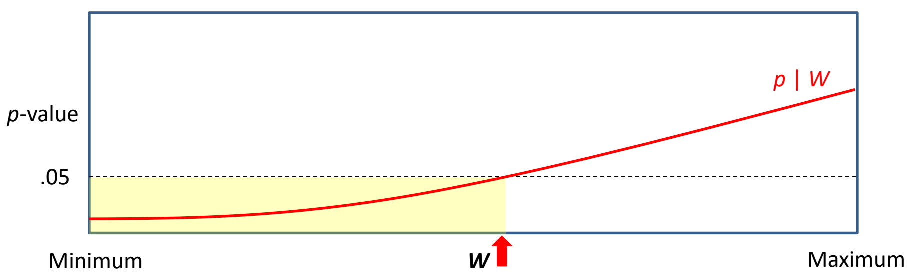
- If one value, this defines a single point of transition between a statistically significant and a statistically nonsignificant conditional effect of the focal predictor, such that $p \le .05$ for either values of the moderator (1) equal to above *W* or (2) equal to and below *W*.

- If two values, this defines the two points of transition between a statistically significant and a statistically nonsignificant conditional effect of the focal predictor, such that the conditional effect is statistically significant for either (1) values of the moderator between the two values of *W*, or (2) values of the moderator at least as large as the larger *W* and at least as small as the smaller *W*.

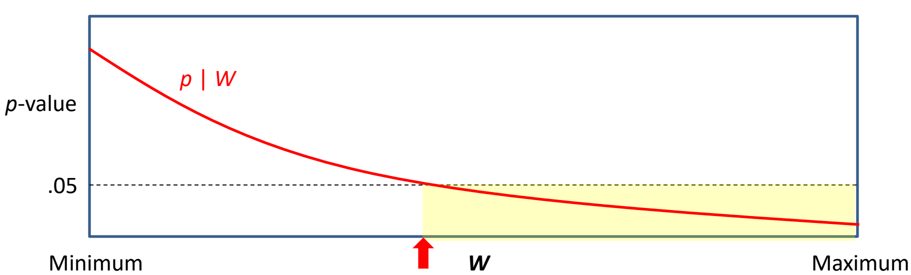
- If no values, that means the conditional effect is statistically significant for ALL values of the moderator within the range of the data, or it NEVER is.

We would not attempt to do this by hand

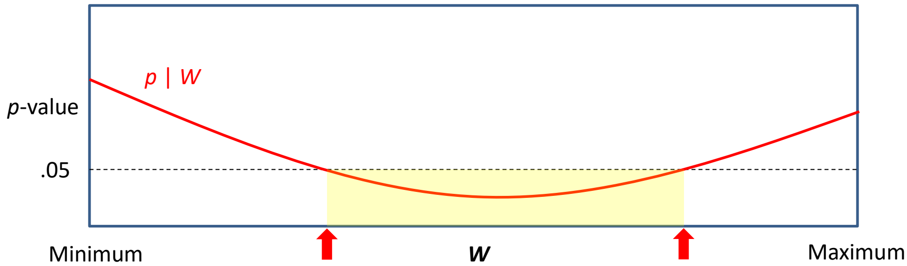
Examples of one solution



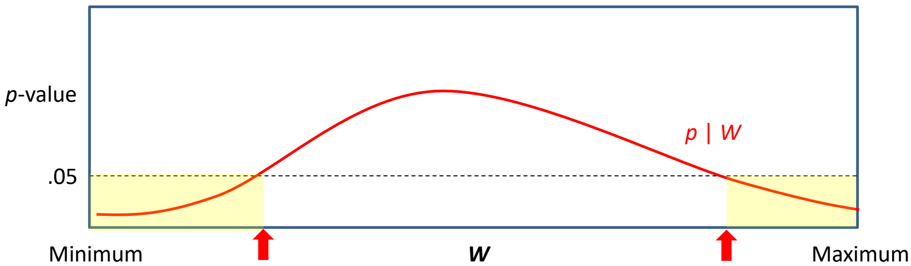
OR



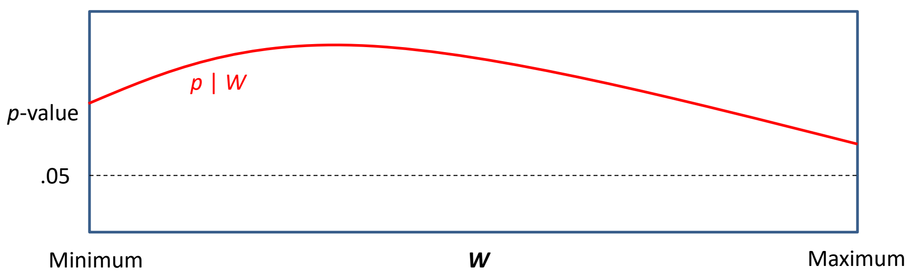
Examples of two solutions



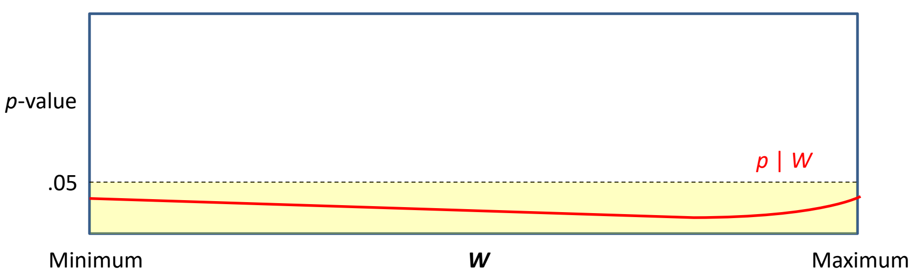
OR



Examples of no solutions



OR



Johnson-Neyman output from PROCESS

```
process cov = treathrs crave0/y=crave2/x=mbrp/w=bdi0
  /model=1/jn=1/plot=1/moments=1.
```

```
%process (data=mbrp,cov=treathrs crave0,y=crave2,x=mbrp,
   w=bdi0,model=1,jn=1,plot=1,moments=1);
```

Output H

```
Moderator value(s) defining Johnson-Neyman significance region(s)
     Value    % below    % above
     .9681    21.4286    78.5714

Conditional effect of X on Y at values of the moderator:
      bdi0     Effect         se          t          p       LLCI       ULCI
     .0000      .5872      .5241     1.1204      .2642     -.4478     1.6222
     .1070      .4858      .4806     1.0108      .3136     -.4632     1.4347
     .2140      .3843      .4373      .8787      .3809     -.4793     1.2479
     .3210      .2828      .3946      .7167      .4746     -.4964     1.0620
     .4280      .1813      .3525      .5144      .6077     -.5147      .8773
     .5350      .0798      .3112      .2565      .7979     -.5348      .6944
     .6420     -.0217      .2713     -.0798      .9365     -.5574      .5141
     .7490     -.1231      .2334     -.5276      .5985     -.5840      .3377
     .8560     -.2246      .1986    -1.1312      .2596     -.6167      .1675
     .9630     -.3261      .1688    -1.9318      .0551     -.6595      .0072
     .9681     -.3309      .1676    -1.9747      .0500     -.6618      .0000
    1.0700     -.4276      .1472    -2.9047      .0042     -.7183     -.1369
    1.1770     -.5291      .1377    -3.8435      .0002     -.8009     -.2573
    1.2840     -.6306      .1426    -4.4220      .0000     -.9122     -.3490
    1.3910     -.7321      .1607    -4.5553      .0000    -1.0494     -.4147
    1.4980     -.8335      .1882    -4.4288      .0000    -1.2052     -.4619
    1.6050     -.9350      .2216    -4.2186      .0000    -1.3727     -.4973
    1.7120    -1.0365      .2587    -4.0063      .0001    -1.5474     -.5256
    1.8190    -1.1380      .2981    -3.8178      .0002    -1.7266     -.5494
    1.9260    -1.2395      .3389    -3.6571      .0003    -1.9088     -.5702
    2.0330    -1.3410      .3808    -3.5216      .0006    -2.0929     -.5890
    2.1400    -1.4424      .4234    -3.4072      .0008    -2.2784     -.6064
```

$\theta_{X \rightarrow Y} | W$

W

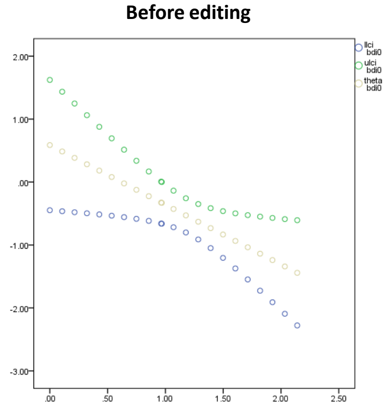
78.6% of the data are up here

A visual representation (SPSS)

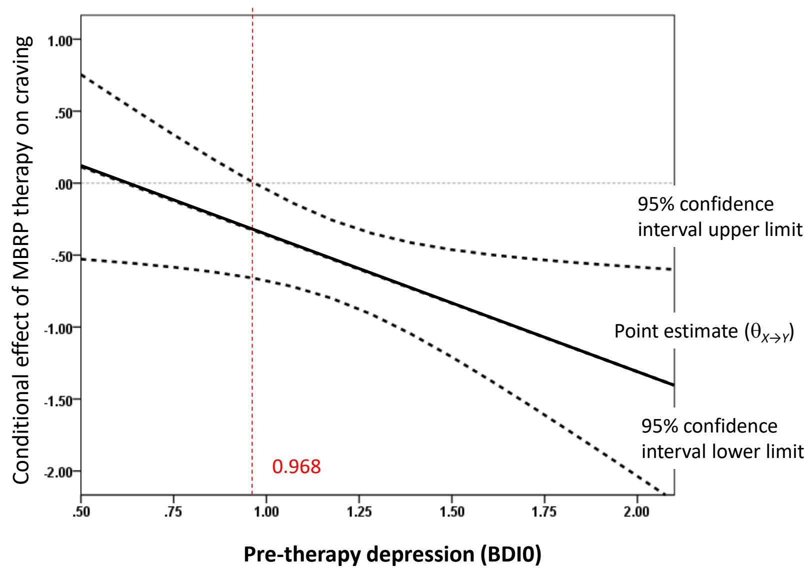
bdi0 effect LLCI ULCI (from PROCESS JN output)

```
data list free/bdi0 theta llci ulci.
begin data.
 .0000      .5872      -.4478     1.6222
 .1070      .4858      -.4632     1.4347
 .2140      .3843      -.4793     1.2479
 .3210      .2828      -.4964     1.0620
 .4280      .1813      -.5147      .8773
 .5350      .0798      -.5348      .6944
 .6420     -.0217      -.5574      .5141
 .7490     -.1231      -.5840      .3377
 .8560     -.2246      -.6167      .1675
 .9630     -.3261      -.6595      .0072
 .9681     -.3309      -.6618      .0000
1.0700     -.4276      -.7183     -.1369
1.1770     -.5291      -.8009     -.2573
1.2840     -.6306      -.9122     -.3490
1.3910     -.7321     -1.0494     -.4147
1.4980     -.8335     -1.2052     -.4619
1.6050     -.9350     -1.3727     -.4973
1.7120    -1.0365     -1.5474     -.5256
1.8190    -1.1380     -1.7266     -.5494
1.9260    -1.2395     -1.9088     -.5702
2.0330    -1.3410     -2.0929     -.5890
2.1400    -1.4424     -2.2784     -.6064
end data.
graph
  /scatterplot(overlay)=bdi0 bdi0 bdi0 WITH llci ulci theta (pair).
```

Before editing

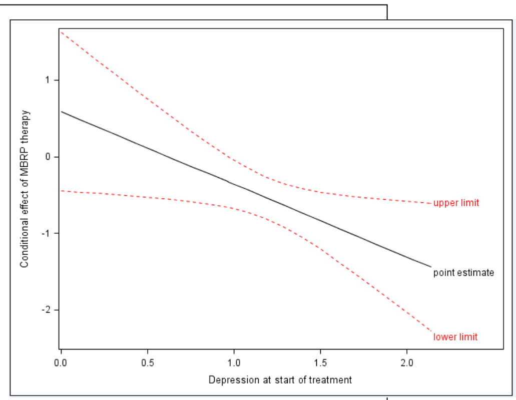


After some editing in SPSS



A visual representation (SAS)

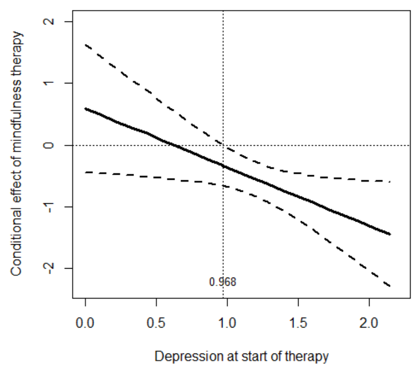
```
data;
input bdi0 effect llci ulci;
datalines;
 .0000      .5872     -.4478    1.6222
 .1070      .4858     -.4632    1.4347
 .2140      .3843     -.4793    1.2479
 .3210      .2828     -.4964    1.0620
 .4280      .1813     -.5147     .8773
 .5350      .0798     -.5348     .6944
 .6420     -.0217     -.5574     .5141
 .7490     -.1231     -.5840     .3377
 .8560     -.2246     -.6167     .1675
 .9630     -.3261     -.6595     .0072
 .9681     -.3309     -.6618     .0000
1.0700     -.4276     -.7183    -.1369
1.1770     -.5291     -.8009    -.2573
1.2840     -.6306     -.9122    -.3490
1.3910     -.7321    -1.0494    -.4147
1.4980     -.8335    -1.2052    -.4619
1.6050     -.9350    -1.3727    -.4973
1.7120    -1.0365    -1.5474    -.5256
1.8190    -1.1380    -1.7266    -.5494
1.9260    -1.2395    -1.9088    -.5702
2.0330    -1.3410    -2.0929    -.5890
2.1400    -1.4424    -2.2784    -.6064
run;
proc sgplot;
  series x=bdi0 y=ulci/curvelabel = "upper limit" lineattrs=(color=red pattern=ShortDash);
  series x=bdi0 y=effect/curvelabel = "point estimate" lineattrs=(color=black pattern=Solid);
  series x=bdi0 y=llci/curvelabel = "lower limit" lineattrs=(color=red pattern=ShortDash);
  yaxis label = "Conditional effect of MBRP therapy";
  xaxis label = "Depression at start of treatment";
run;
```



A visual representation (R)

```
bdi0<-c(0,.107,.214,.321,.438,.535,.642,.749,.856,.963,.968,1.070,1.177,
  1.284,1.391,1.498,1.605,1.712,1.819,1.926,2.033,2.140)
effect<-c(.587,.486,.384,.283,.181,.080,-.022,-.123,-.225,-.326,-.331,-.428,
  -.529,-.631,-.732,-.834,-.935,-1.037,-1.138,-1.240,-1.341,-1.442)
llci<-c(-.448,-.463,-.479,-.496,-.515,-.535,-.557,-.584,-.617,-.660,-.662,
  -.718,-.801,-.912,-1.049,-1.205,-1.373,-1.547,-1.727,-1.909,-2.092,-2.278)
ulci<-c(1.622,1.435,1.248,1.062,.877,.684,.515,.338,.168,.007,0,-.137,
  -.257,-.349,-.415,-.462,-.497,-.526,-.549,-.571,-.589,-.606)
plot(x=bdi0,y=effect,type="l",pch=19,ylim=c(-2.3,2),xlim=c(0,2.2),lwd=3,
ylab="Conditional effect of mindfulness therapy",
xlab="Depression at start of therapy")
points(bdi0,llci,lwd=2,lty=2,type="l")
points(bdi0,ulci,lwd=2,lty=2,type="l")
abline(h=0,untf = FALSE,lty=3,lwd=1)
abline(v=0.968,untf=FALSE,lty=3,lwd=1)
text(0.968,-2.2,"0.968",cex=0.8)
```

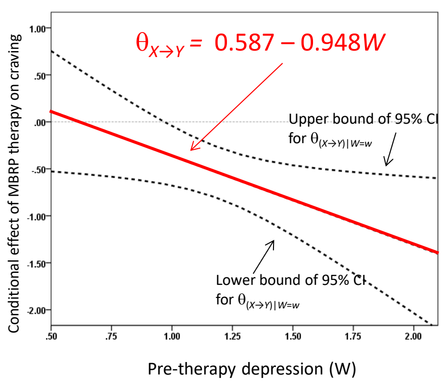
From the JN option in PROCESS.



Testing moderation versus testing a conditional effect

A test of moderation is a test as to whether the size of the effect of *X* on *Y* is related to the proposed moderator. This is different than asking whether a conditional effect is different from zero.

$$\hat{Y} = 1.038 + 0.587X + 1.122W - 0.948XW + \ldots$$



$\theta_{X\to Y} = b_1 + b_3M = 0.587 - 0.948W$

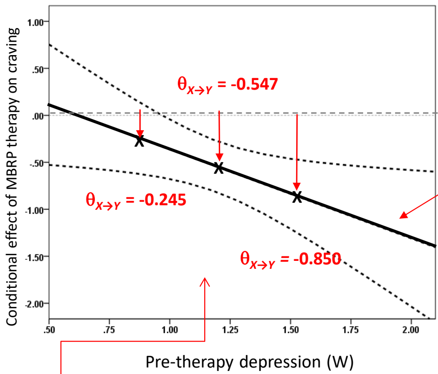
b_3 is the slope of this line. It is statistically different from zero, meaning that the effect of *X* depends on *W*---moderation.

Moderation does **not** imply that the conditional effect of *X* is different from zero at some, any, or all specific values of the moderator that you choose. Often it will be, perhaps for some values of the moderator but not others. But this is not a requirement of moderation.

Testing moderation versus testing a conditional effect

A test of moderation is a test as to whether the size of the effect of *X* on *Y* is related to the proposed moderator. This is different than asking whether a conditional effect is different from zero.

When testing a conditional effect, we are asking whether the effect of *X* on Y at a specific value of *W* is statistically different from zero. This is the difference between the point estimate of $\theta_{X \rightarrow Y}$ and zero.



$$\theta_{X \rightarrow Y} = b_1 + b_3W = 0.587 - 0.948W$$

Difference in significance does not imply significantly different. The pattern of significance or lack thereof across values of *M* does not say anything about moderation.

Conditional effect of X on Y at values of the moderator(s):

bdi0	Effect	se	t	p	LLCI	ULCI
.8772	-.2447	.1922	-1.2733	.2047	-.6243	.1348
1.1963	-.5473	.1375	-3.9818	.0001	-.8188	-.2759
1.5153	-.8500	.1933	-4.3973	.0000	-1.2317	-.4683

Comparing conditional effects

We want to know whether the conditional effect of *X* on *Y* when $W = w_1$ is different from the conditional effect of *X* on *Y* when $W = w_2$.

$$\begin{aligned}\theta_{(X \rightarrow Y)\,|\,W=w_2} - \theta_{(X \rightarrow Y)\,|\,W=w_1} &= (b_1 + b_3w_2) - (b_1 + b_3w_1) \\ &= (w_2 - w_1)b_3\end{aligned}$$

and the standard error of the difference is $(w_2 - w_1) \times$ standard error of b_3. Under the null hypothesis that the difference in conditional effects is zero, the ratio

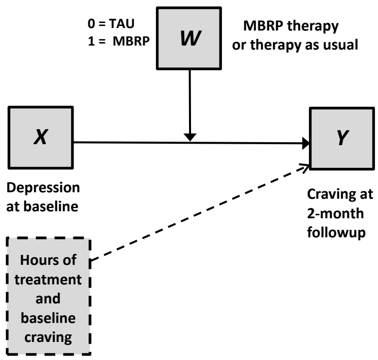
$$\frac{(w_2 - w_1)b_3}{(w_2 - w_1)se_{b_3}}$$

is distributed as $t(df_{residual})$. But notice that *regardless of the values of* w_1 *and* w_2, this ratio simplifies to b_3 / se_{b_3}. **We already have the *p*-value for this**. It is the *p*-value for b_3 from the regression model.

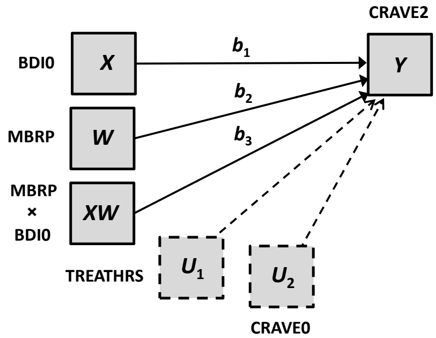
A test of linear moderation of *X*'s effect on *Y* by *W* is equivalent to a test of the difference between *any* two conditional effects of *X*. Moderation = any two conditional effects of *X* are different from each other. No moderation = no two conditional effects of *X* are different from each other. It doesn't matter what values of w_1 and w_2 you choose.

A Dichotomous Moderator

Conceptual diagram



Statistical model



Does the association between pre-treatment depression and later craving differ between those who receive MBRP therapy versus therapy as usual?

This model has already been estimated

```
compute mbrpdep = mbrp*bdi0.
regression/dep = crave2/method = enter mbrp bdi0 mbrpdep treathrs crave0.
```

```
data mbrp;set mbrp;mbrpdep=mbrp*bdi0;run;
proc reg data=mbrp;model crave2=mbrp bdi0 mbrpdep treathrs crave0;run;
```

X = BDI0
W = MBRP
Y = CRAVE2

$$\hat{Y} = 1.038 + 1.122X + 0.587W - 0.948XW + \ldots$$

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.038	.470		2.209	.029
	MBRP: Therapy as usual (0) or MBRP therapy (1)	.587	.524	.299	1.120	.264
	BDI0: Beck Depression Inventory baseline	1.122	.276	.366	4.063	.000
	mbrpdep	-.948	.423	-.598	-2.240	.026
	TREATHRS: Hours of therapy	-.018	.010	-.120	-1.719	.088
	CRAVE0: Baseline craving	.192	.073	.183	2.614	.010

a. Dependent Variable: CRAVE2: Craving at two month follow-up

b_1 = 1.122
b_2 = 0.587
b_3 = -0.948

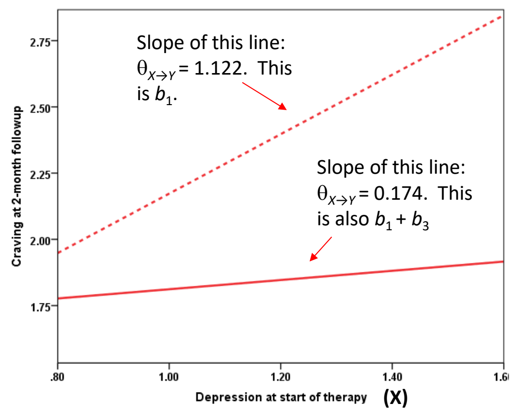
The coefficient for the product is statistically different from zero. This means that the effect of pre-treatment depression on later craving differs between those given MBRP therapy and those given therapy as usual.

A graphical depiction of the model

$\hat{Y} = 1.038 + 1.122X + 0.587W - 0.948XW + ...$ or, equivalently,

$\hat{Y} = 1.038 + (1.122 - 0.948W)X + 0.587W + ...$

Therapy as usual (**MBRP** (*W*) =0)
MBRP therapy (**MBRP** (*W*) = 1)



The conditional effect of pre-therapy depression ($\theta_{X\rightarrow Y}$) is defined by the function

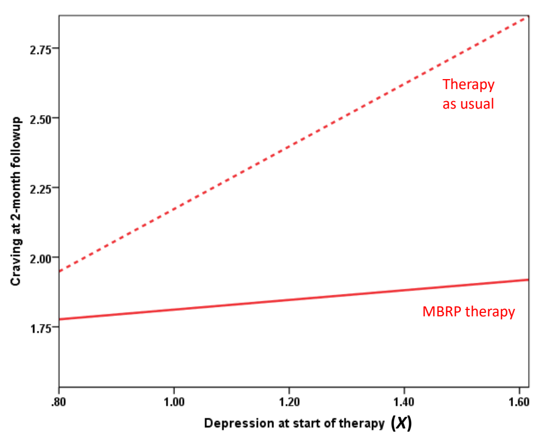
$$\theta_{X\rightarrow Y} = b_1 + b_3W$$
$$= 1.122 - 0.948W$$

MBRP (*W*)	$\theta_{X\rightarrow Y}$
0	1.122
1	0.174

Two cases that differ by one unit on *W* are estimated to differ by $\theta_{X\rightarrow Y}$ units on *Y*. $\theta_{X\rightarrow Y}$ depends on *W*.

Substantive interpretation of the pattern

Therapy as usual (**MBRP** (*W*) =0)
MBRP therapy (**MBRP** (*W*) = 1)



A larger effect of pre-therapy depression on later craving among those who experienced therapy as usual compared to those who received mindfulness behavioral relapse prevention therapy. MBRP therapy seems to have disrupted the link between depression and craving.

Probing the Interaction

When the moderator is dichotomous, the pick-a-point procedure is the only option available, as the Johnson-Neyman technique is meaningful only with a quantitative moderator. Typically, you'd want to estimate the effect of the focal predictor at the two values of the moderator and conduct an inferential test for each conditional effect.

$$\hat{Y} = 1.038 + 1.122X + 0.587W - 0.948XW$$
$$= 1.038 + (1.122 - 0.948W)X + 0.587W + \ldots$$
$$\theta_{X \to Y} = b_1 + b_3 W = 1.122 - 0.948W$$

When one of the moderator categories is coded 0, we already have an estimate of $\theta_{X \to Y}$ when $W = 0$. That estimate is b_1. And the regression output provides a test of significance.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.038	.470		2.209	.029
	MBRP: Therapy as usual (0) or MBRP therapy (1)	.587	.524	.299	1.120	.264
	BDI0: Beck Depression Inventory baseline	1.122	.276	.366	4.063	.000
	mbrpdep	-.948	.423	-.598	-2.240	.026
	TREATHRS: Hours of therapy	-.018	.010	-.120	-1.719	.088
	CRAVE0: Baseline craving	.192	.073	.183	2.614	.010

a. Dependent Variable: CRAVE2: Craving at two month follow-up

Among those given therapy as usual, those who were relatively more depressed at the start of therapy had relatively higher craving at two months follow-up, $\theta_{X \to Y} = 1.122$, $t(162) = 4.063$, $p < .001$

Probing the Interaction

We already know effect of pre-therapy depression on later craving among those given MBRP therapy. That is 1.122 – 0.948(1) = 0.174. We can use the regression centering approach, constructing $W' = W - 1$ and reestimating the model in order to get a test of significance.

```
compute mbrp_p = mbrp-1.
compute mbrpdep = mbrp_p*bdi0.
regression/dep = crave2/method = enter mbrp_p bdi0 mbrpdep treathrs crave0.
```

```
data mbrp;set mbrp;mbrp_p=mbrp-1;mbrpdep=mbrp_p*bdi0;
proc reg data=mbrp;model crave2=mbrp_p bdi0 mbrpdep treathrs crave0;run;
```

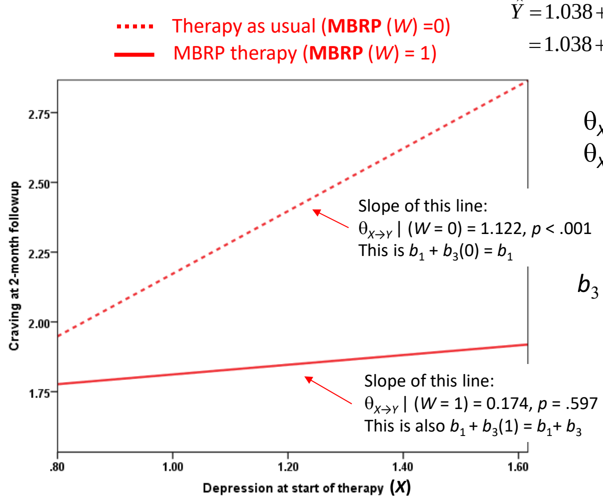
Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.626	.535		3.041	.003
	mbrp_p	.587	.524	.299	1.120	.264
	BDI0: Beck Depression Inventory baseline	.174	.328	.057	.529	.597
	mbrpdep	-.948	.423	-.639	-2.240	.026
	TREATHRS: Hours of therapy	-.018	.010	-.120	-1.719	.088
	CRAVE0: Baseline craving	.192	.073	.183	2.614	.010

a. Dependent Variable: CRAVE2: Craving at two month follow-up

Among those given MBRP therapy, there was no statistically significant relationship between pre-therapy depression and later craving, $\theta_{X \to Y} = 0.174$, $t(162) = 0.529$, $p = .597$

Interpreting b_3



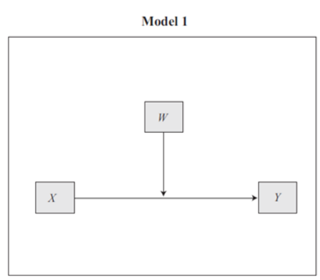
$$\hat{Y} = 1.038 + 1.122X + 0.587W - 0.948XW$$
$$= 1.038 + (1.122 - 0.948W)X + 0.587W + \ldots$$

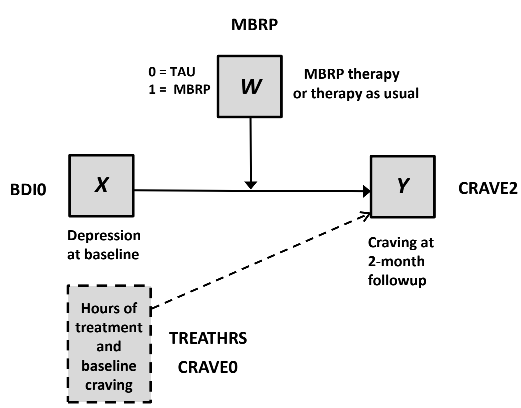
$$\theta_{X \to Y} | W = b_1 + b_3 W$$
$$\theta_{X \to Y} | W = 1.122 - 0.948W$$

$$\begin{aligned} b_3 &= \theta_{X \to Y} | (W = 1) - \theta_{X \to Y} | (W = 0) \\ &= (b_1 + b_3) - (b_1) \\ &= (0.174) - (1.122) \\ &= -0.948 \end{aligned}$$

So b_3 is the difference in the slopes of these two lines. As W increases by one unit, $\theta_{X \to Y}$ decreases by 0.948 units. This difference is statistically different from zero.

Estimation Using PROCESS






```
process cov = treathrs crave0/y=crave2/x=bdi0/w=mbrp/model=1.
```

```
%process (data=mbrp,cov=treathrs crave0,y=crave2,x=bdi0,w=mbrp,model=1);
```

PROCESS Output

Output I

```
Model = 1
    Y = crave2
    X = bdi0
    M = mbrp

Statistical Controls:
CONTROL= treathrs crave0

Sample size
        168

**************************************************************************
Outcome: crave2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5140      .2642      .7277    11.6319     5.0000   162.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.0385      .4701     2.2090      .0286      .1102     1.9668
mbrp          .5872      .5241     1.1204      .2642     -.4478     1.6222
bdi0         1.1221      .2762     4.0625      .0001      .5767     1.6675
int_1        -.9485      .4235    -2.2398      .0265    -1.7847     -.1122
treathrs     -.0177      .0103    -1.7190      .0875     -.0380      .0026
crave0        .1920      .0735     2.6138      .0098      .0470      .3371

Interactions:

 int_1    bdi0        X      mbrp

R-square increase due to interaction(s):
         R2-chng          F        df1        df2          p
int_1      .0228     5.0166     1.0000   162.0000      .0265

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

Conditional effect of X on Y at values of the moderator(s):
       mbrp     Effect         se          t          p       LLCI       ULCI
      .0000     1.1221      .2762     4.0625      .0001      .5767     1.6675
     1.0000      .1736      .3281      .5291      .5974     -.4744      .8216
```

PROCESS detects that the moderator is dichotomous and generates the conditional effect of the focal predictor at the two values of the moderator.

Moderation analysis summary

- A moderator of the effect of *X* on *Y* is a variable which influences or otherwise is related to the size of *X*'s effect on *Y*.

- Including a variable defined as the product *X* of *W* to a regression model that includes *X* and *W* allows *X*'s effect on *Y* to be a linear function of *W*.

- The regression coefficient for *XW* in such a model is hard to interpret without a picture. Draw a picture of your model before attempting to interpret.

- We can dissect or "probe" interactions in a few different ways:

  - The pick-a-point approach requires us to select values of *W* at which to estimate the conditional effect of *X* on *Y*. Usually the selection is arbitrary.
  - The Johnson-Neyman technique avoids the need to choose values of the moderator arbitrarily.

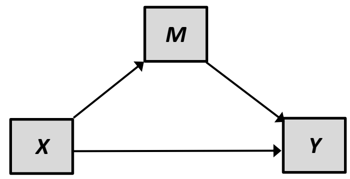
- Care must be taken when interpreting the regression coefficients for *X* and *W* in a model that includes *XW*. They are not "main effects" and they may not have any substantive interpretation. Their interpretation will be influenced by their scaling and whether a value of zero is meaningful on the measurement scale. We can make it meaningful by centering.

Combining moderation and mediation
"Conditional Process Analysis"

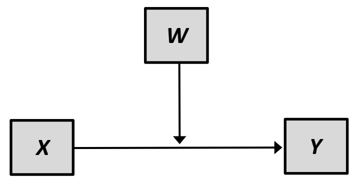
"Conditional process analysis" is a general modeling strategy undertaken with the goal of describing the *conditional* nature of the *mechanism*(*s*) by which a variable transmits its effect on another, and testing hypotheses about such contingent effects.

A merging of two ideas conceptually and analytically:

"**Process analysis**", used to quantify and examine the direct and indirect pathways through which an antecedent variable *X* transmits its effect on a consequent variable *Y* through an intermediary *M*. Better known as "mediation analysis" these days.



"**Moderation analysis**" used to examine how the effect of an antecedent *X* on an consequent *Y* depends on a third moderator variable *W* (a.k.a. "interaction")



History

Idea is not new (e.g., Judd & Kenny, 1981; James & Brett, 1984; Baron and Kenny, 1986). It goes by various names that often confuse, including "moderated mediation" and "mediated moderation."

<u>**More recently:**</u>

Muller, Judd, and Yzerbyt (2005): Describe analytical models and steps for assessing when "mediation is moderated" and "moderation is mediated."

Edwards and Lambert (2007): Take a path analysis perspective and show how various effects in a simple mediation model can be conditioned on a third variable.

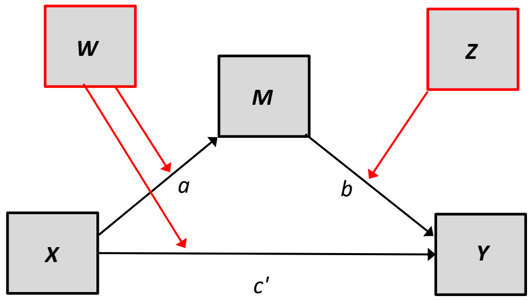
Preacher, Rucker, and Hayes (2007): Provide a formal definition of the *conditional indirect effect* and give formulas, standard errors, and a bootstrap approach for estimating and testing hypotheses about moderated mediation in five different models.

MacKinnon and colleagues (e.g., Fairchild & MacKinnon, 2009): Explicate various analytical approaches to testing hypotheses about mediated moderation and moderated mediation.

Hayes (2013). Introduces the term "conditional process modeling" (also see Hayes and Preacher, 2013) and provides tools for SPSS and SAS to make it easy to do.

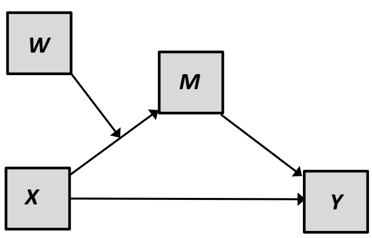
Hayes (2015): Introduces the *index of moderated mediation* which provides a formal test for moderated mediation in a variety of models.

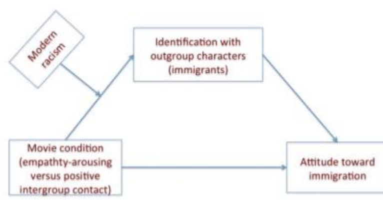
"Moderated mediation"



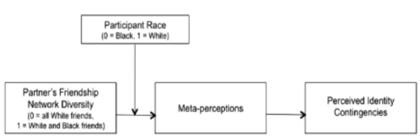
- The indirect effect of *X* on *Y* through *M* is estimated as the product of the *a* and *b* paths
- But what if the size of *a* or *b* (or both) depends on another variable (i.e., is moderated)?
- If so, then the magnitude of the indirect effect therefore depends on a third variable, meaning that "mediation is moderated".
- When *a* or *b* is moderated, it is sensible then to estimate "conditional indirect effects"—values of indirect effect conditioned on values of the moderator variable that moderates *a* and/or *b*.
- Direct effects can also be conditional. For instance, above, *W* moderates *X*'s direct effect on *Y*.

Examples: *X* to *M* path moderated by *W*

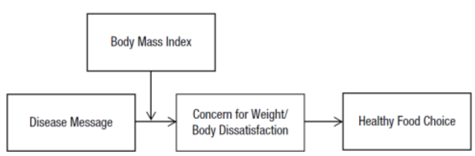




Igartua, J-J., & Frutos, F. J. (2017). Enhancing attitudes toward stigmatized groups with movies. Mediating and moderating processes of narrative persuasion. *International Journal of Communication, 11*, 158-`77.

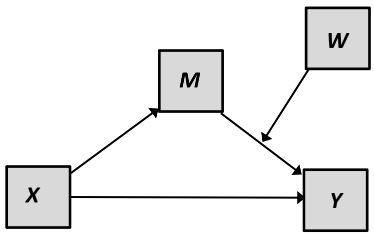


Wout, D. A., Murphy, M. C., & Steele, C. M. (2010). When your friends matter: The effect of White students' racial friendship networks on meta-perceptions and Perceived identity contingencies. ***Journal of Experimental Social Psychology****, 46*, 1035-1041.

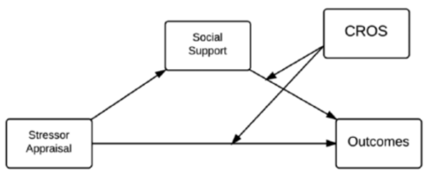


Hoyt, C. L., Burnette, J. L., & Auster-Gussman, L. (2014). "Obesity is a disease": Examining the self-regulatory impact of this public-health message. ***Psychological Science****, 25*, 997-1002.

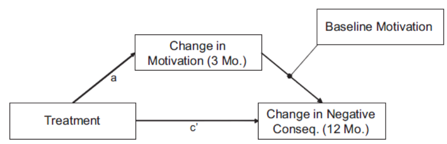
Examples: *M* to *Y* path moderated by *W*



Cornelissen, G., Bashshur, M. R., Rode, J., & Le Menestrel, M. (2013). Rules or consequences? The role of ethical mind-sets in moral dynamics. ***Psychological Science, 24***, 492-488.

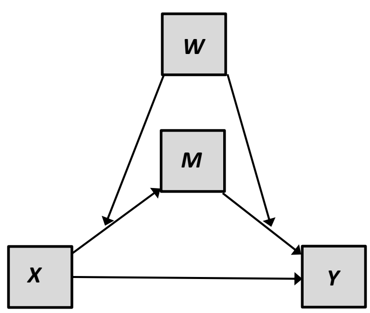


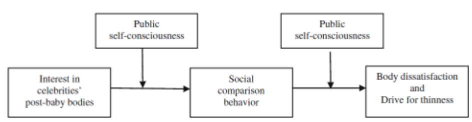
Boren, J. P., & Veksler, A. E. (2015). Communicatively restricted organizational stress (CROS) I: Conceptualization and overview. ***Management Communication Quarterly,*** *29*, 28-55.



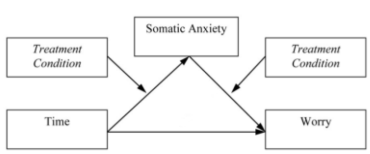
Stein, L. A. R., Minugh, P. A. et al. (2009). Readiness to change as a mediator of the effect of a brief motivational intervention on posttreatment alcohol-related consequences of injured emergency department hazardous drinkers. ***Psychology of Addictive Behaviors, 23***, 185-195.

Examples: *X* to *M* and *M* to *Y* path moderated by *W*

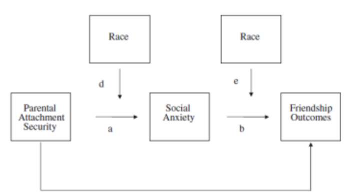




Chae, J. (2014). Interest in celebrities' post-baby bodies and Korean women's body image disturbance after childhood. ***Sex Roles,*** *71*, 419-435.



Donegan, E., & Dugas, M. (2012). Generalized anxiety disorder: A comparison of symptom change in adults receiving cognitive-behavioral therapy or applied relaxation. ***Journal of Consulting and Clinical Psychology,*** *80*, 490-496.



Parade, S. H., Leerkes, E. M., & Blankson, A. (2010). Attachment to parents, social anxiety, and close relationships of female students over the transition to college**.** ***Journal of Youth and Adolescence***, *39*, 127-137.

“Conditional direct effect”

In a mediation model, the direct effect of *X* on *Y* quantifies *X*’s effect independent of the intervening variable or variables. If that direct effect is moderated, then the direct effect is conditional on the variable that moderates *X*’s effect. For example,

$$\hat{M} = i_1 + aX$$

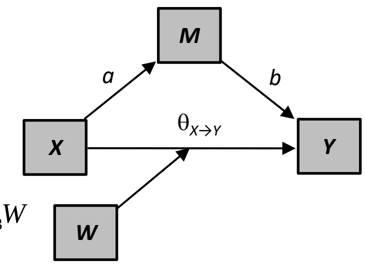
$$\hat{Y} = i_2 + c_1'X + c_2'W + c_3'XW + bM$$

or, equivalently,

$$\hat{Y} = i_2 + (c_1' + c_3'W)X + c_2'W + bM$$

or, equivalently,

$$\hat{Y} = i_2 + \theta_{X\rightarrow Y}X + c_2'W + bM \text{ where } \theta_{X\rightarrow Y} = c_1' + c_3W$$



In this model, $\theta_{X\rightarrow Y}$ is the conditional direct effect of *X*, which is defined by the function $c'_1 + c'_3W$. Holding *M* constant, two cases that differ by one unit on *X* are estimated to differ by $c'_1 + c'_3W$ units on *Y*.

This is a very basic conditional process model. It models two pathways through which *X* affects *Y*. One is unconditional and indirect via *M*, and the other is direct but conditional—the size of the direct effect depends on *W*.

“Conditional direct effect”

In a mediation model, the direct effect of *X* on *Y* quantifies *X*’s effect independent of the intervening variable or variables. If that direct effect is moderated, then the direct effect is conditional on the variable that moderates *X*’s effect. For example,

$$\boxed{\hat{M} = i_1 + aX}$$

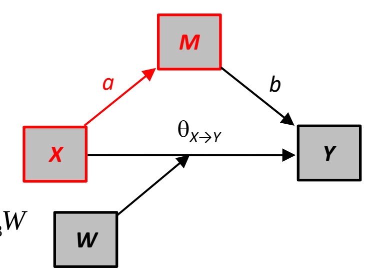
$$\hat{Y} = i_2 + c_1'X + c_2'W + c_3'XW + bM$$

or, equivalently,

$$\hat{Y} = i_2 + (c_1' + c_3'W)X + c_2'W + bM$$

or, equivalently,

$$\hat{Y} = i_2 + \theta_{X\rightarrow Y}X + c_2'W + bM \text{ where } \theta_{X\rightarrow Y} = c_1' + c_3W$$



In this model, $\theta_{X\rightarrow Y}$ is the conditional direct effect of *X*, which is defined by the function $c'_1 + c'_3W$. Holding *M* constant, two cases that differ by one unit on *X* are estimated to differ by $c'_1 + c'_3W$ units on *Y*.

This is a very basic conditional process model. It models two pathways through which *X* affects *Y*. One is unconditional and indirect via *M*, and the other is direct but conditional—the size of the direct effect depends on *W*.

“Conditional direct effect”

In a mediation model, the direct effect of *X* on *Y* quantifies *X*’s effect independent of the intervening variable or variables. If that direct effect is moderated, then the direct effect is conditional on the variable that moderates *X*’s effect. For example,

$$\hat{M} = i_1 + aX$$

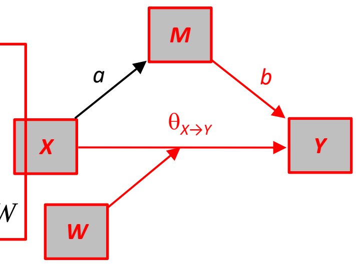
$$\hat{Y} = i_2 + c_1'X + c_2'W + c_3'XW + bM$$

or, equivalently,

$$\hat{Y} = i_2 + (c_1' + c_3'W)X + c_2'W + bM$$

or, equivalently,

$$\hat{Y} = i_2 + \theta_{X \to Y}X + c_2'W + bM \text{ where } \theta_{X \to Y} = c_1' + c_3W$$



In this model, $\theta_{X \to Y}$ is the conditional direct effect of *X*, which is defined by the function $c'_1 + c'_3W$. Holding *M* constant, two cases that differ by one unit on *X* are estimated to differ by $c'_1 + c'_3W$ units on *Y*.

This is a very basic conditional process model. It models two pathways through which *X* affects *Y*. One is unconditional and indirect via *M*, and the other is direct but conditional—the size of the direct effect depends on *W*.

“Conditional indirect effect”

An indirect effect is quantified as the product of paths linking *X* to *Y* via the intermediary variable. If one of those paths depends on a moderator, then so too does the indirect effect depend on that moderator. For example:

$$\hat{M} = i_1 + aX$$

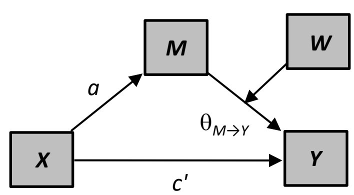
$$\hat{Y} = i_2 + c'X + b_1M + b_2W + b_3MW$$

or, equivalently,

$$\hat{Y} = i_2 + c'X + (b_1 + b_3W)M + b_2W$$

or, equivalently,

$$\hat{Y} = i_2 + c'X + \theta_{M \to Y}M + b_2W \text{ where } \theta_{M \to Y} = b_1 + b_3W$$



The indirect effect of *X* on *Y* via *M* is $a\theta_{M \to Y}$, but as $\theta_{M \to Y}$ is a conditional effect (the conditional effect of *M*), then $a\theta_{M \to Y}$ is the *conditional indirect effect* of *X* on *Y* via *M*: $a\theta_{M \to Y} = a(b_1 + b_3W) = ab_1 + ab_3W$. It depends on *W*.

This is also a basic conditional process model, and potentially more interesting one. It allows for the process or ‘mechanism’ linking *X* to *Y* via *M* to differ systematically as a function of *W*. This model allows “mediation to be moderated.”

"Conditional indirect effect"

An indirect effect is quantified as the product of paths linking *X* to *Y* via the intermediary variable. If one of those paths depends on a moderator, then so too does the indirect effect depend on that moderator. For example:

$$\boxed{\hat{M} = i_1 + aX}$$

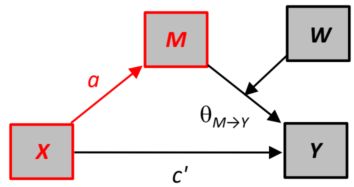
$$\hat{Y} = i_2 + c'X + b_1M + b_2W + b_3MW$$

or, equivalently,

$$\hat{Y} = i_2 + c'X + (b_1 + b_3W)M + b_2W$$

or, equivalently,

$$\hat{Y} = i_2 + c'X + \theta_{M \to Y}M + b_2W \text{ where } \theta_{M \to Y} = b_1 + b_3W$$



The indirect effect of *X* on *Y* via *M* is $a\theta_{M \to Y}$, but as $\theta_{M \to Y}$ is a conditional effect (the conditional effect of *M*), then $a\theta_{M \to Y}$ is the *conditional indirect effect* of *X* on *Y* via *M*: $a\theta_{M \to Y} = a(b_1 + b_3W) = ab_1 + ab_3W$. It depends on *W*.

This is also a basic conditional process model, and potentially more interesting one. It allows for the process or 'mechanism' linking *X* to *Y* via *M* to differ systematically as a function of *W*. This model allows "mediation to be moderated."

"Conditional indirect effect"

An indirect effect is quantified as the product of paths linking *X* to *Y* via the intermediary variable. If one of those paths depends on a moderator, then so too does the indirect effect depend on that moderator. For example:

$$\hat{M} = i_1 + aX$$

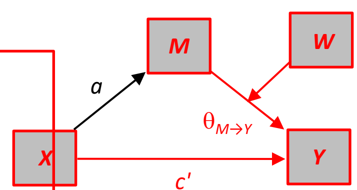
$$\hat{Y} = i_2 + c'X + b_1M + b_2W + b_3MW$$

or, equivalently,

$$\hat{Y} = i_2 + c'X + (b_1 + b_3W)M + b_2W$$

or, equivalently,

$$\hat{Y} = i_2 + c'X + \theta_{M \to Y}M + b_2W \text{ where } \theta_{M \to Y} = b_1 + b_3W$$

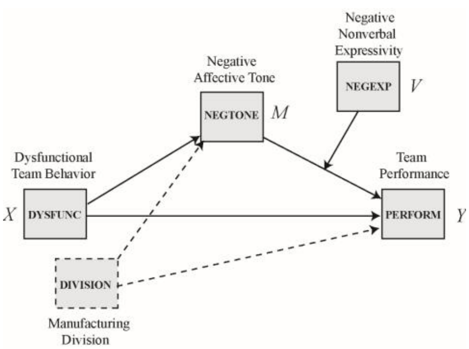


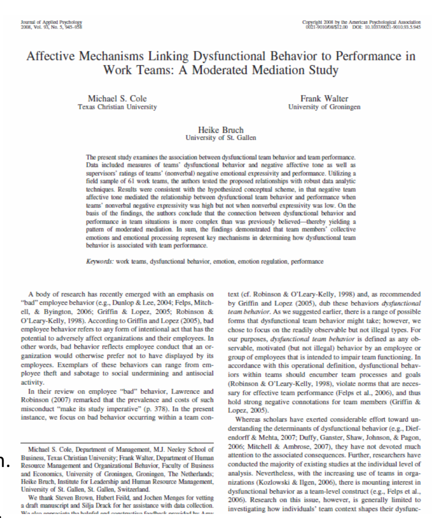
The indirect effect of *X* on *Y* via *M* is $a\theta_{M \to Y}$, but as $\theta_{M \to Y}$ is a conditional effect (the conditional effect of *M*), then $a\theta_{M \to Y}$ is the *conditional indirect effect* of *X* on *Y* via *M*: $a\theta_{M \to Y} = a(b_1 + b_3W) = ab_1 + ab_3W$. It depends on *W*.

This is also a basic conditional process model, and potentially more interesting one. It allows for the process or 'mechanism' linking *X* to *Y* via *M* to differ systematically as a function of *W*. This model allows "mediation to be moderated."

Condition process analysis: Example 1

Condition process analysis: Example 1





Journal of Applied Psychology
2008, Vol. 93, No. 5, 945–958

0021-9010/08/$12.00 DOI: 10.1037/0021-9010.93.5.945

Affective Mechanisms Linking Dysfunctional Behavior to Performance in Work Teams: A Moderated Mediation Study

Michael S. Cole
Texas Christian University

Frank Walter
University of Groningen

Heike Bruch
University of St. Gallen

The present study examines the association between dysfunctional team behavior and team performance. Data included measures of teams' dysfunctional behavior and negative affective tone as well as supervisors' ratings of teams' (nonverbal) negative emotional expressivity and performance. Utilizing a field sample of 61 work teams, the authors tested the proposed relationships with robust data analytic techniques. Results were consistent with the hypothesized conceptual scheme, in that negative team affective tone mediated the relationship between dysfunctional team behavior and performance when teams' nonverbal negative expressivity was high but not when nonverbal expressivity was low. On the basis of the findings, the authors conclude that the connection between dysfunctional behavior and performance in team situations is more complex than was previously believed—thereby yielding a pattern of moderated mediation. In sum, the findings demonstrated that team members' collective emotions and emotional processing represent key mechanisms in determining how dysfunctional team behavior is associated with team performance.

Keywords: work teams, dysfunctional behavior, emotion, emotion regulation, performance

A body of research has recently emerged with an emphasis on "bad" employee behavior (e.g., Dunlop & Lee, 2004; Felps, Mitchell, & Byington, 2006; Griffin & Lopez, 2005; Robinson & O'Leary-Kelly, 1998). According to Griffin and Lopez (2005), bad employee behavior refers to any form of intentional act that has the potential to adversely affect organizations and their employees. In other words, bad behavior reflects employee conduct that an organization would otherwise prefer not to have displayed by its employees. Exemplars of these behaviors can range from employee theft and sabotage to social undermining and antisocial activity.

In their review on employee "bad" behavior, Lawrence and Robinson (2007) remarked that the prevalence and costs of such misconduct "make its study imperative" (p. 378). In the present instance, we focus on bad behavior occurring within a team context (cf. Robinson & O'Leary-Kelly, 1998) and, as recommended by Griffin and Lopez (2005), dub these behaviors *dysfunctional team behavior*. As we suggested earlier, there is a range of possible forms that dysfunctional team behavior might take; however, we chose to focus on the readily observable but not illegal types. For our purposes, *dysfunctional team behavior* is defined as any observable, motivated (but not illegal) behavior by an employee or group of employees that is intended to impair team functioning. In accordance with this operational definition, dysfunctional behaviors within teams should encumber team processes and goals (Robinson & O'Leary-Kelly, 1998), violate norms that are necessary for effective team performance (Felps et al., 2006), and thus hold strong negative connotations for team members (Griffin & Lopez, 2005).

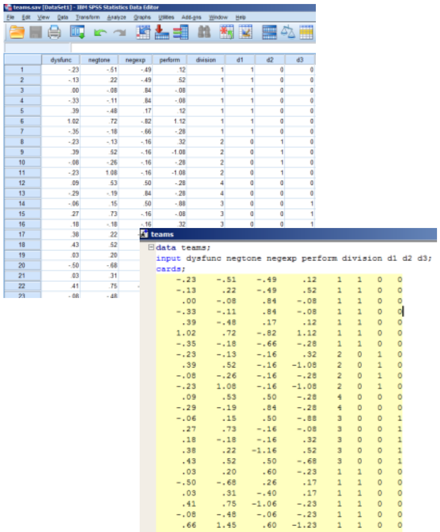
Whereas scholars have exerted considerable effort toward understanding the determinants of dysfunctional behavior (e.g., Diefendorff & Mehta, 2007; Duffy, Ganster, Shaw, Johnson, & Pagon, 2006; Mitchell & Ambrose, 2007), they have not devoted much attention to the associated consequences. Further, researchers have conducted the majority of existing studies at the individual level of analysis. Nevertheless, with the increasing use of teams in organizations (Kozlowski & Ilgen, 2006), there is mounting interest in dysfunctional behavior as a team-level construct (e.g., Felps et al., 2006). Research on this issue, however, is generally limited to investigating how individuals' team context shapes their dysfunc-

Michael S. Cole, Department of Management, M.J. Neeley School of Business, Texas Christian University; Frank Walter, Department of Human Resource Management and Organizational Behavior, Faculty of Business and Economics, University of Groningen, Groningen, The Netherlands; Heike Bruch, Institute for Leadership and Human Resource Management, University of St. Gallen, St. Gallen, Switzerland.

We thank Steven Brown, Hubert Feild, and Jochen Menges for vetting a draft manuscript and Silja Drack for her assistance with data collection.

This is a model of **negative affective tone (*M*)** as the mechanism by which **dysfunctional team behavior (*X*)** influences **performance *(Y)*,** with that mechanism being contingent on the extent to which **team members hide their negative feelings (*V*)** from the team. This "nonverbal expressivity" is postulated as moderating the effect of negative tone on performance. This is a "second stage" moderated mediation model.

The Data: TEAMS



```
data teams;
input dysfunc negtone negexp perform division d1 d2 d3;
cards;
   -.23   -.51   -.49    .12   1   1   0   0
   -.13    .22   -.49    .52   1   1   0   0
    .00   -.08    .84   -.08   1   1   0   0
   -.33   -.11    .84   -.08   1   1   0   0
    .39   -.48    .17    .12   1   1   0   0
   1.02    .72   -.82   1.12   1   1   0   0
   -.35   -.18   -.66   -.28   1   1   0   0
   -.23   -.13   -.16    .32   2   0   1   0
    .39    .52   -.16  -1.08   2   0   1   0
   -.08   -.26   -.16   -.28   2   0   1   0
   -.23   1.08   -.16  -1.08   2   0   1   0
    .09    .53    .50   -.28   4   0   0   0
   -.29   -.19    .84   -.28   4   0   0   0
   -.06    .15    .50   -.88   3   0   0   1
    .27    .73   -.16   -.08   3   0   0   1
    .18   -.18   -.16    .32   3   0   0   1
    .38    .22  -1.16    .52   3   0   0   1
    .43    .52    .50   -.68   3   0   0   1
    .03    .20    .60   -.23   1   1   0   0
   -.50   -.68    .26    .17   1   1   0   0
    .03    .31   -.40    .17   1   1   0   0
    .41    .75  -1.06   -.23   1   1   0   0
   -.08   -.48   -.06   -.23   1   1   0   0
    .66   1.45    .60  -1.23   1   1   0   0
```

60 teams working in an automobile parts manufacturing facility.

DYSFUNC: Dysfunctional team behavior, i.e., How often members of the team do things to weaken the work of others or hinder change and innovation.

NEGTONE: Negative affective group tone. How often team members report feeling negative emotions at work such as "angry", "disgust", etc.

NEGEXP: Negative nonverbal expressivity. Supervisor's perception as to how easy it is to tell how team members are feeling.

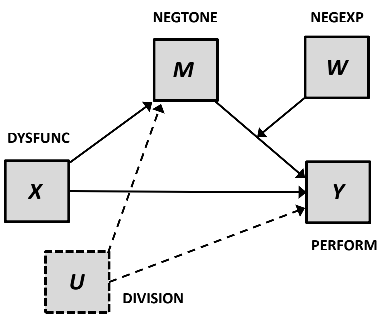
PERFORM: Team performance. Supervisor's judgment as to the team's efficiency, ability to get task done in a timely fashion, etc.

All variables are scaled arbitrarily, but higher = "more"

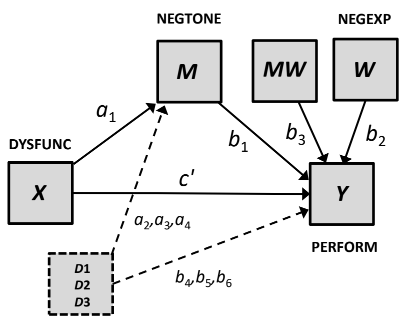
Also available is which of four parts divisions the team worked in, as a single categorical variable (division) as well as three dummy variables (d1, d2, d3).

Conceptual and statistical models

Conceptual Model



Statistical Model



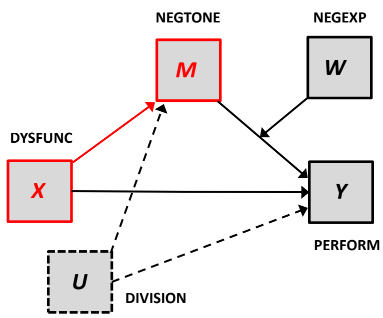
$$\hat{M} = i_1 + a_1 X + a_2 D_1 + a_3 D_2 + a_4 D_3$$

$$\hat{Y} = i_1 + c'X + b_1 M + b_2 W + b_3 MW + b_4 D_1 + b_5 D_2 + b_6 D_3$$

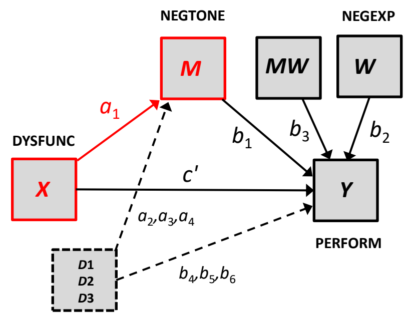
The number of equations needed is equal to the number of variables with arrows pointing **at** them in the conceptual or statistical diagram.

Conceptual and statistical models

Conceptual Model



Statistical Model



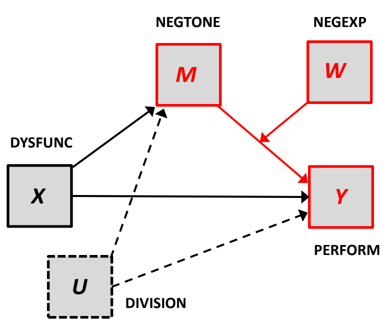
$$\hat{M} = i_1 \boxed{+ a_1 X} + a_2 D_1 + a_3 D_2 + a_4 D_3$$

$$\hat{Y} = i_1 + c'X + b_1 M + b_2 W + b_3 MW + b_4 D_1 + b_5 D_2 + b_6 D_3$$

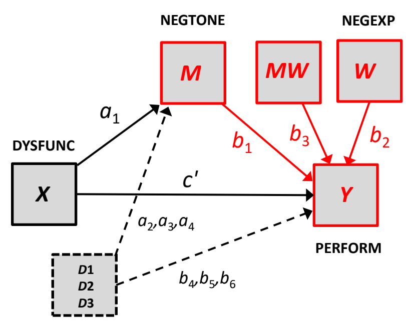
The effect of dysfunctional team behavior (*X*) on negative affective tone of the work environment (*M*).

Conceptual and statistical models

Conceptual Model



Statistical Model



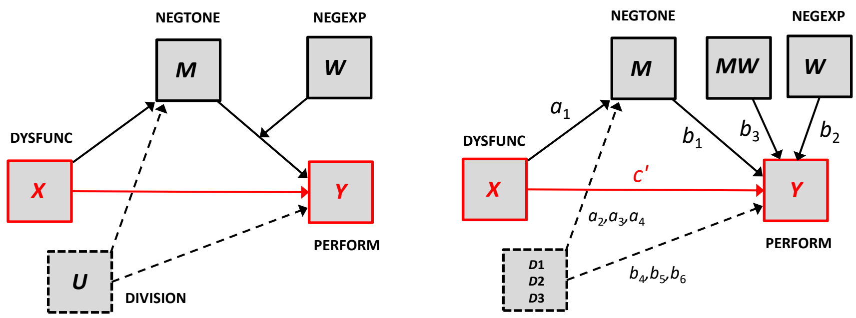
$$\hat{M} = i_1 + a_1 X + a_2 D_1 + a_3 D_2 + a_4 D_3$$

$$\hat{Y} = i_1 + c'X \boxed{+ b_1 M + b_2 W + b_3 MW} + b_4 D_1 + b_5 D_2 + b_6 D_3$$

The moderation of the effect of the negative affective tone of the work environment (*M*) on team performance (*Y*) by negative nonverbal expressivity (*W*).

Conceptual and statistical models

Conceptual Model | Statistical Model



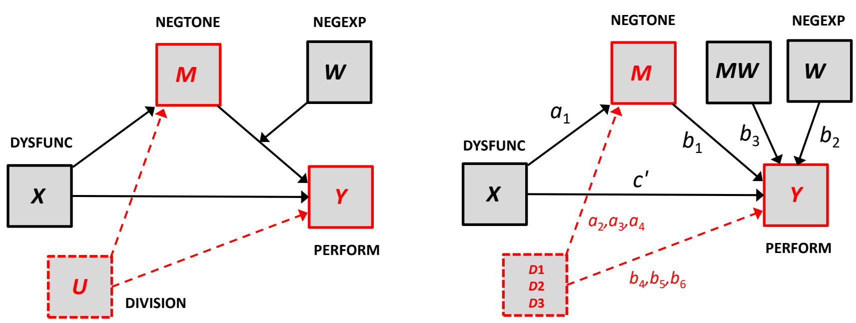
$$\hat{M} = i_1 + a_1X + a_2D_1 + a_3D_2 + a_4D_3$$

$$\hat{Y} = i_1 \boxed{+ c'X} + b_1M + b_2W + b_3MW + b_4D_1 + b_5D_2 + b_6D_3$$

The direct effect of dysfunctional team behavior (*X*) on team performance (*Y*).

Conceptual and statistical models

Conceptual Model | Statistical Model



$$\hat{M} = i_1 + a_1X \boxed{+ a_2D_1 + a_3D_2 + a_4D_3}$$

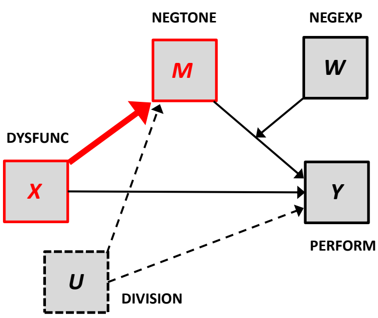
$$\hat{Y} = i_1 + c'X + b_1M + b_2W + b_3MW \boxed{+ b_4D_1 + b_5D_2 + b_6D_3}$$

Covariates to account for potential confounding by divisional differences (*U*) in negative tone of the work environment (*M*) and performance (*Y*)

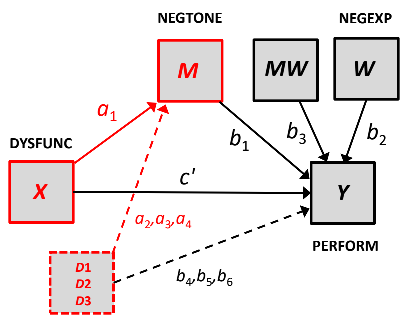
Estimating the a_1 path

Let's first estimate the effect of dysfunctional team behavior on the negative affective tone of the team environment: Path a_1 in the statistical model.

Conceptual Model



Statistical Model



$$\hat{M} = i_1 + a_1 X + a_2 D_1 + a_3 D_2 + a_4 D_3$$

Emphasis in not on statistical significance, as neither the direct or indirect effects of *X* are defined entirely in terms of a_1.

Estimating the a_1 path

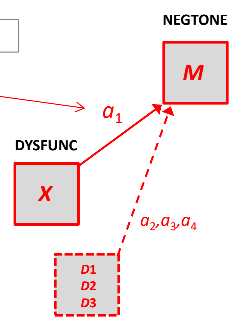
```
regression/dep=negtone/method=enter dysfunc d1 d2 d3.
```

```
proc reg data=teams;model negtone=dysfunc d1 d2 d3;run;
```

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-.206	.130		-1.576	.121
	Dysfunctional team behavior	.609	.167	.431	3.655	.001
	d1	.349	.171	.307	2.033	.047
	d2	.295	.212	.193	1.391	.170
	d3	.251	.166	.230	1.508	.137

a. Dependent Variable: Negative affective tone



a_1 = 0.609

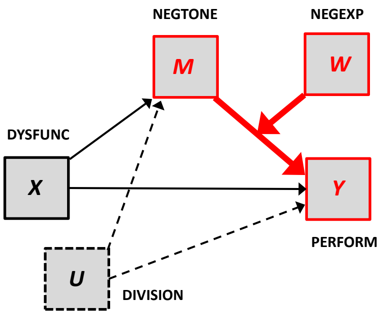
$$\hat{M} = -0.206 + 0.609X + 0.349D_1 + 0.295D_2 + 0.251D_3$$

Teams whose members exhibit relatively more dysfunctional behavior tend to operate in a work environment characterized by relatively more negative affective tone (i.e., members report more negative affect)

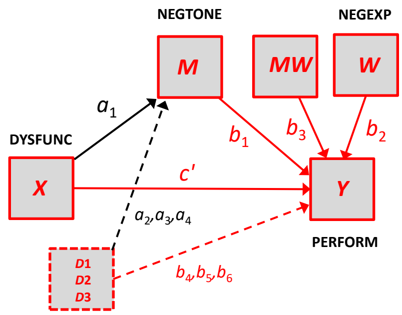
Estimating the moderation component of the model

The conceptual model proposes that the effect of negative work tone on performance depends on negative nonverbal expressivity. Let's see whether there is evidence of this.

Conceptual Model



Statistical Model



$$\hat{Y} = i_1 + c'X + \boxed{b_1M + b_2W + b_3MW} + b_4D_1 + b_5D_2 + b_6D_3$$

We most care about the moderation components of the model of *Y*: b_1, b_2, and b_3. But these must be estimated in the context of the complete model of *Y*, which includes *X* as well.

Estimating the moderation component of the model

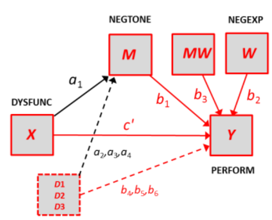
```
compute toneexp=negtone*negexp.
regression/dep=perform/method=enter dysfunc negtone negexp toneexp d1 d2 d3.
```

```
data teams;set teams;toneexp=negtone*negexp;run;
proc reg data=teams;model perform=dysfunc negtone negexp toneexp d1 d2 d3;run;
```

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-.175	.130		-1.344	.185
	Dysfunctional team behavior	.373	.181	.265	2.062	.044
	Negative affective tone	-.489	.138	-.491	-3.549	.001
	Negative expressivity	-.022	.118	-.023	-.188	.852
	toneexp	-.450	.245	-.240	-1.835	.072
	d1	.182	.172	.161	1.056	.296
	d2	.084	.210	.055	.400	.690
	d3	.282	.165	.259	1.709	.093

a. Dependent Variable: Team performance



b_1 = -0.489, b_2=-0.022, b_3 = -0.450

$$\hat{Y} = -0.175 + 0.373X - 0.489M - 0.022W - 0.450MW + 0.182D_1 + 0.084D_2 + 0.282D_3$$

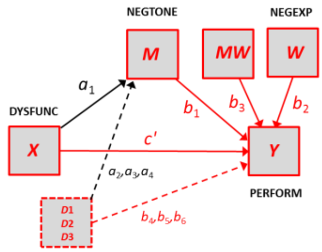
"Marginally significant" evidence that the effect of negative tone of the work environment on team performance depends on the negative nonverbal expressivity of team members. To better understand this, dissect this model.

Estimating the moderation component of the model

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-.175	.130		-1.344	.185
	Dysfunctional team behavior	.373	.181	.265	2.062	.044
	Negative affective tone	-.489	.138	-.491	-3.549	.001
	Negative expressivity	-.022	.118	-.023	-.188	.852
	toneexp	-.450	.245	-.240	-1.835	.072
	d1	.182	.172	.161	1.056	.296
	d2	.084	.210	.055	.400	.690
	d3	.282	.165	.259	1.709	.093

a. Dependent Variable: Team performance



b_1 = 0.489, b_3 = -0.450

$$\hat{Y} = -0.175 + 0.373X - 0.489M - 0.022W - 0.450MW + 0.182D_1 + 0.084D_2 + 0.282D_3$$

which can be written as

$$\hat{Y} = -0.175 + 0.373X + (-0.489 - 0.450W)M - 0.022W + 0.182D_1 + 0.084D_2 + 0.282D_3$$

or

$$\hat{Y} = -0.175 + 0.373X + \theta_{M \to Y}M - 0.022W + 0.182D_1 + 0.084D_2 + 0.282D_3$$

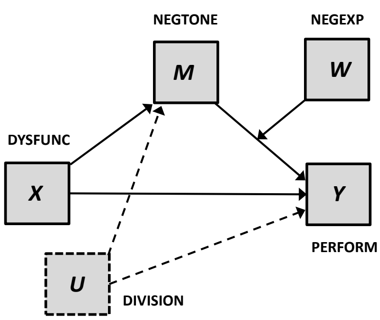
where

$$\theta_{M \to Y} = -0.489 - 0.450W = b_1 + b_3W$$

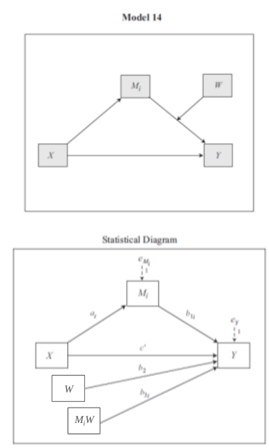
Let's visualize and probe this. PROCESS will take the work out of it.

Estimation of the model in PROCESS

PROCESS takes most of the computational burden off our shoulders.



This is PROCESS model 14




```
process cov = d1 d2 d3/x=dysfunc/m=negtone/y=perform/w=negexp/boot=10000/model=14/plot
= 1.
```

```
%process (data=teams,cov=d1 d2 d3,x=dysfunc,m=negtone,y=perform,w=negexp,
          boot=10000,model=14, plot = 1);
```

PROCESS output

Output J

```
OUTCOME VARIABLE:
 perform

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5937      .3524      .2006     4.0428     7.0000    52.0000      .0013

Model
              coeff         se          t          p       LLCI       ULCI
constant     -.1754      .1305    -1.3444      .1847     -.4373      .0864
dysfunc       .3729      .1808     2.0622      .0442      .0100      .7357
negtone      -.4886      .1377    -3.5485      .0008     -.7649     -.2123
negexp       -.0221      .1176     -.1875      .8520     -.2581      .2140
Int_1        -.4498      .2451    -1.8353      .0722     -.9417      .0420
d1            .1815      .1720     1.0556      .2960     -.1635      .5266
d2            .0841      .2099      .4004      .6905     -.3372      .5053
d3            .2816      .1648     1.7087      .0935     -.0491      .6123

Product terms key:
 Int_1    :        negtone   x        negexp

Test(s) of highest order unconditional interaction(s):
       R2-chng          F        df1        df2          p
M*W      .0419     3.3684     1.0000    52.0000      .0722
```

b_2 = -0.022
b_1 = -0.489
b_3 = -0.450

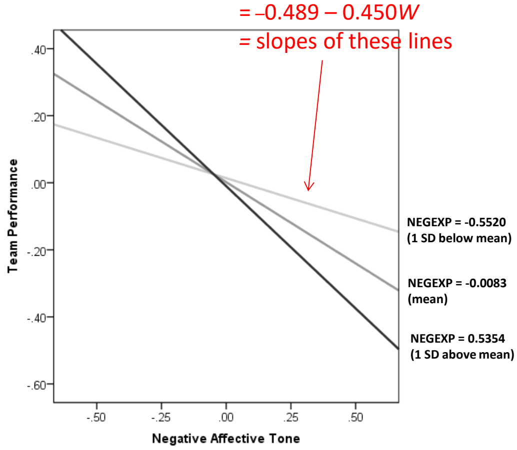
Visualizing the interaction

Use of the PLOT option in SPSS (**`plot=1`**) produces a table of estimated values of the outcome for various combinations of the focal predictor and moderator, and an SPSS program to generate a skeleton of the plot that can be edited.

M *W* $\hat{Y}$

```
DATA LIST FREE/negtone negexp perform.
BEGIN DATA.
     -.4782     -.5520      .1288
      .0472     -.5520      .0026
      .5726     -.5520     -.1237
     -.4782     -.0083      .2338
      .0472     -.0083     -.0210
      .5726     -.0083     -.2757
     -.4782      .5354      .3387
      .0472      .5354     -.0445
      .5726      .5354     -.4278
END DATA.
GRAPH/SCATTERPLOT=negtone WITH perform
   BY negexp.
```

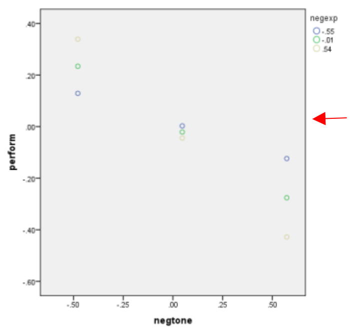
$\theta_{M \rightarrow Y} = b_1 + b_3W$ =
$= -0.489 - 0.450W$
= slopes of these lines



Visualizing the interaction

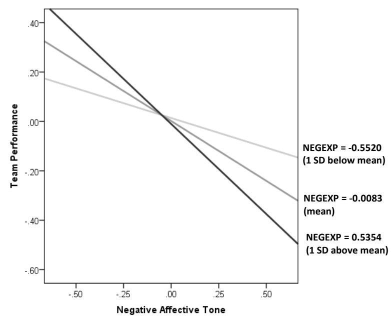
```
DATA LIST FREE/negtone negexp perform.
BEGIN DATA.
      -.4782     -.5520      .1288
       .0472     -.5520      .0026
       .5726     -.5520     -.1237
      -.4782     -.0083      .2338
       .0472     -.0083     -.0210
       .5726     -.0083     -.2757
      -.4782      .5354      .3387
       .0472      .5354     -.0445
       .5726      .5354     -.4278
END DATA.
GRAPH/SCATTERPLOT=negtone WITH perform BY negexp.
```

Use of the PLOT option (**`plot=1`**) produces a table of estimated values of the outcome for various combinations of the focal predictor and moderator.



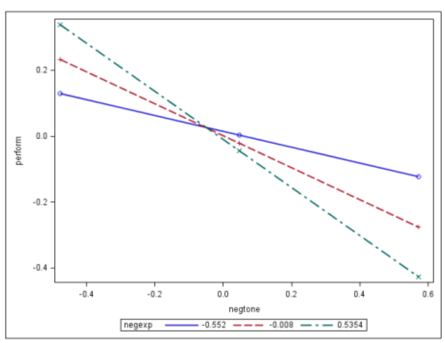
Before editing

After editing



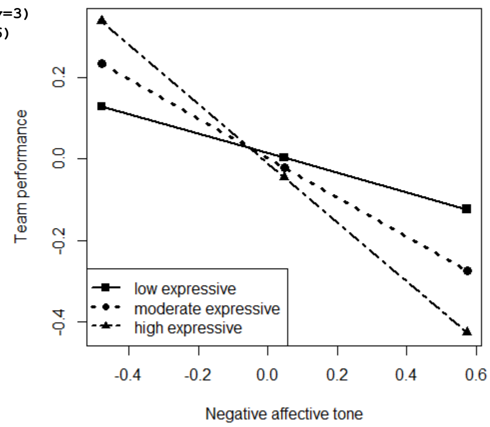
Example code in SAS

```
data;
input negtone negexp perform;
cards;
-.4782     -.5520      .1288
 .0472     -.5520      .0026
 .5726     -.5520     -.1237
-.4782     -.0083      .2338
 .0472     -.0083     -.0210
 .5726     -.0083     -.2757
-.4782      .5354      .3387
 .0472      .5354     -.0445
 .5726      .5354     -.4278
run;
proc sgplot;reg x=negtone y=perform/group=negexp;run;
```



Example code in R

```
m<-c(-.478,.047,.573,-.478,.047,.573,-.478,.047,.573)
w<-c(-.552,-.552,-.552,-.008,-.008,-.008,.535,.535,.535)
y<-c(.129,.003,-.124,.234,-.021,-.276,.339,-.045,-.428)
wmarker<-c(15,15,15,16,16,16,17,17,17)
plot(y=y,x=m,cex=1.2,pch=wmarker,xlab="Negative affective tone",
ylab="Team performance")
legend.txt<-c("low expressive","moderate expressive","high expressive")
legend("bottomleft", legend = legend.txt,cex=1,lty=c(1,3,6),lwd=c(2,3,2),pch=c(15,16,17))
lines(m[w==-.552],y[w==-.552],lwd=2)
lines(m[w==-.008],y[w==-.008],lwd=3,lty=3)
lines(m[w==.535],y[w==.535],lwd=2,lty=6)
```



Probing the interaction: Pick-a-point

PROCESS sees that the moderator is quantitative (i.e., it has more than 2 values) so it implements the pick-a-point procedure with moderator values equal to 14th, 50th, and 84th percentile.

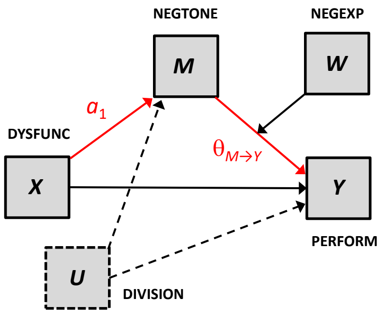
$$\theta_{M \rightarrow Y} = b_1 + b_3 W = -0.489 - 0.450W$$

```
*************************************************************************
 Focal predict: negtone  (M)
       Mod var: negexp   (W)

Conditional effects of the focal predictor at values of the moderator(s):

    negexp     Effect        se         t          p       LLCI       ULCI
    -.5308     -.2498     .2196   -1.1379      .2604     -.6904      .1907
    -.0600     -.4616     .1434   -3.2188      .0022     -.7494     -.1738
     .6000     -.7585     .1633   -4.6451      .0000    -1.0862     -.4308
```

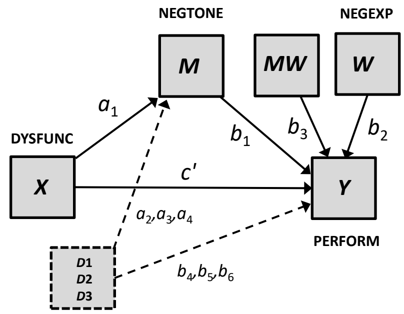
Negative affective tone is significantly negatively related to performance among teams relatively “moderate” and “relatively high” in negative nonverbal expressivity but not among teams “relatively low” in negative nonverbal expressivity.

The conditional indirect effect of *X*

Conceptual Model



Statistical Model

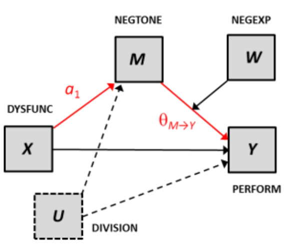


The conditional indirect effect of *X* on *Y* through *M* is the product of the effect of X on *M* (a_1) and the conditional effect of *M* on *Y* given *W* ($\theta_{M \to Y} = b_1 + b_3 W$):

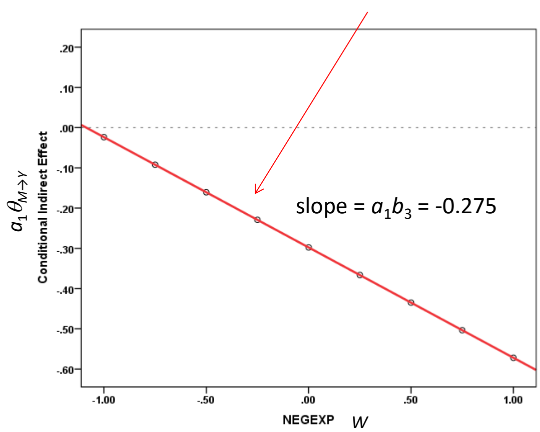
$$a_1 \theta_{M \to Y} = a_1(b_1 + b_3 W) = 0.610(-0.489 - 0.450W)$$

The indirect effect of dysfunctional team behavior on team performance through negative tone is allowed to be a function of negative nonverbal expressivity.

A visual representation of the indirect effect



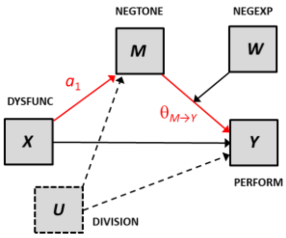
$$a_1 \theta_{M \to Y} = a_1(b_1 + b_3 W) = 0.610(-0.489 - 0.450W)$$
$$= a_1 b_1 + a_1 b_3 W = -0.298 - 0.275W$$



The indirect effect is increasingly negative as negative nonverbal expressivity increases. The "**index of moderated mediation**" is $a_1 b_3$ = -0.275. It quantifies the relationship between the moderator and the indirect effect in this model.

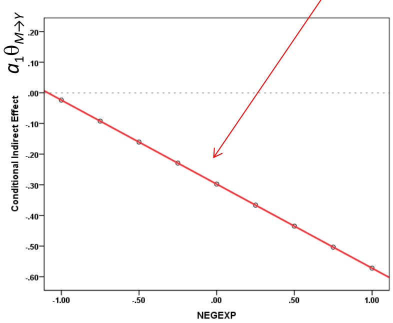
As will be seen, a hypothesis test that the slope of this function is equal to 0 is a formal test of moderated mediation….moderation of the indirect effect.

The conditional indirect effect of *X*



$$a_1\theta_{M\to Y} = a_1(b_1 + b_3W) = 0.610(-0.489 - 0.450W)$$
$$= a_1b_1 + a_1b_3W = -0.298 - 0.275W$$

The indirect effect is increasingly negative as negative nonverbal expressivity increases.



Using this function, we can estimate the indirect effect for any value of the moderator we choose:

If we had used `moments = 1`

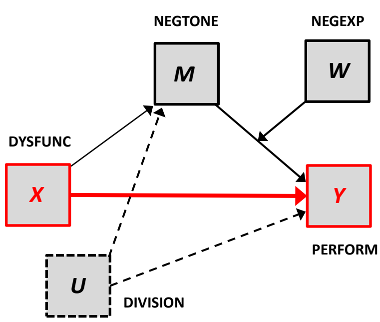
	NEGEXP (*W*)	a_1	$\theta_{M\to Y}$	$a_1\theta_{M\to Y}$
$\bar{W} - SD_W$	-0.552	0.610	-0.240	-0.146
$\bar{W}$	-0.008	0.610	-0.485	-0.296
$\bar{W} + SD_W$	0.535	0.610	-0.730	-0.445

Conditional indirect effects

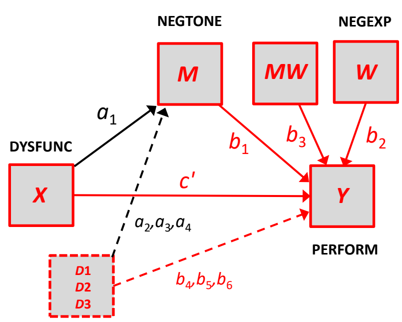
The direct effect of *X*

The direct effect of *X* is the effect of *X* of *Y* that does not operate through *M.*

Conceptual Model



Statistical Model



$$\hat{Y} = i_1 \boxed{+ c'X} + b_1M + b_2W + b_3MW + b_4D_1 + b_5D_2 + b_6D_3$$

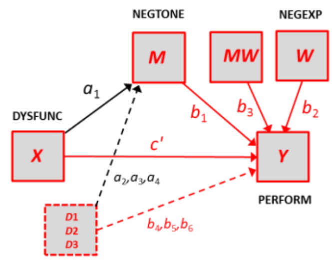
In this model, the direct effect is fixed to be unmoderated. It is a constant rather than a function of another variable in the model.

The direct effect of *X* (we estimated earlier)

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-.175	.130		-1.344	.185
	Dysfunctional team behavior	.373	.181	.265	2.062	.044
	Negative affective tone	-.489	.138	-.491	-3.549	.001
	Negative expressivity	-.022	.118	-.023	-.188	.852
	toneexp	-.450	.245	-.240	-1.835	.072
	d1	.182	.172	.161	1.056	.296
	d2	.084	.210	.055	.400	.690
	d3	.282	.165	.259	1.709	.093

a. Dependent Variable: Team performance



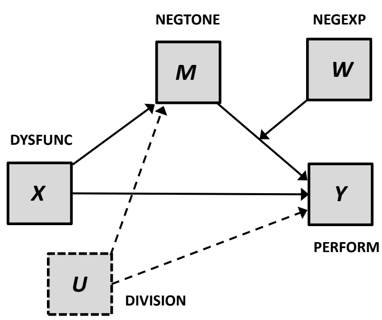
$c' = 0.373$, $t(55) = 2.062$, $p < .05$.

$$\hat{Y} = -0.175 + 0.373X - 0.489M - 0.022W - 0.450MW + 0.182D_1 + 0.084D_2 + 0.282D_3$$

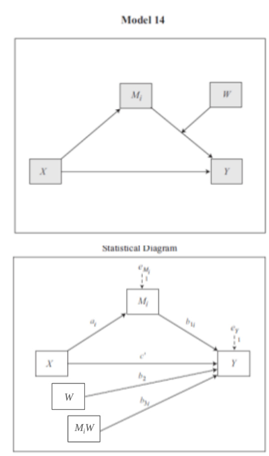
Holding constant negative affective tone and negative nonverbal expressivity, teams that exhibit more dysfunctional behavior perform *better*.

Estimation of the model in PROCESS

PROCESS takes most of the computational burden off our shoulders.



This is PROCESS model 14

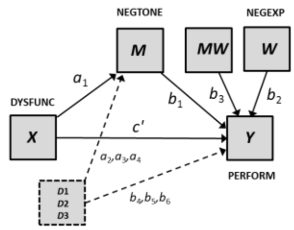



```
process cov = d1 d2 d3/x=dysfunc/m=negtone/y=perform/w=negexp/boot=10000
          /model=14/plot = 1.
```

```
%process (data=teams,cov=d1 d2 d3,x=dysfunc,m=negtone,y=perform,w=negexp,
          boot=10000,model=14, plot = 1);
```

PROCESS output

```
Model = 14
    Y = perform
    X = dysfunc
    M = negtone
    W = negexp

Statistical Controls:
CONTROL= d1       d2       d3

Sample size
         60
```




```
**************************************************************************
Outcome: negtone
```

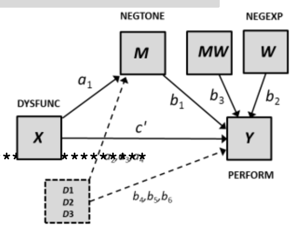
$$\boxed{\hat{M} = -0.206 + 0.610X + \ldots}$$

```
Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5026      .2526      .2213     4.6462     4.0000    55.0000      .0027

Model
              coeff         se          t          p       LLCI       ULCI
constant     -.2057      .1305    -1.5760      .1208     -.4672      .0559
dysfunc       .6095      .1668     3.6546      .0006      .2753      .9437
d1            .3487      .1715     2.0332      .0469      .0050      .6923
d2            .2951      .2122     1.3906      .1700     -.1302      .7204
d3            .2507      .1663     1.5078      .1373     -.0825      .5840

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

$a_1 = 0.610$

Output J

PROCESS output



$$\boxed{\hat{Y} = -0.175 + 0.373X - 0.489M - 0.022W - 0.450MW + \ldots}$$

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

Outcome: perform

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5937      .3524      .2006     4.0428     7.0000    52.0000      .0013

Model
              coeff         se          t          p       LLCI       ULCI
constant     -.1754      .1305    -1.3444      .1847     -.4373      .0864
negtone      -.4886      .1377    -3.5485      .0008     -.7649     -.2123
dysfunc       .3729      .1808     2.0622      .0442      .0100      .7357
negexp       -.0221      .1176     -.1875      .8520     -.2581      .2140
int_1        -.4498      .2451    -1.8353      .0722     -.9417      .0420
d1            .1815      .1720     1.0556      .2960     -.1635      .5266
d2            .0841      .2099      .4004      .6905     -.3372      .5053
d3            .2816      .1648     1.7087      .0935     -.0491      .6123

Interactions:

 int_1    negtone     X     negexp

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

b_1 = -0.489
c' = 0.373
b_2 = -0.022
b_3 = -0.450

Output J

PROCESS output

```
******************** DIRECT AND INDIRECT EFFECTS *************************

Direct effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .3729      .1808     2.0622      .0442      .0100      .7357

INDIRECT EFFECT:
 dysfunc      ->    negtone     ->    perform

     negexp     Effect    BootSE   BootLLCI   BootULCI
     -.5308     -.1523     .1540     -.4335      .1943
     -.0600     -.2813     .1249     -.5432     -.0549
      .6000     -.4623     .1678     -.8095     -.1503

      Index of moderated mediation:
            Index    BootSE   BootLLCI   BootULCI
negexp     -.2742     .1791     -.7172     -.0234

W values in conditional tables are the 16th, 50th, and 84th percentiles.
```

Direct effect
$c' = .373, p < .05$

Output J

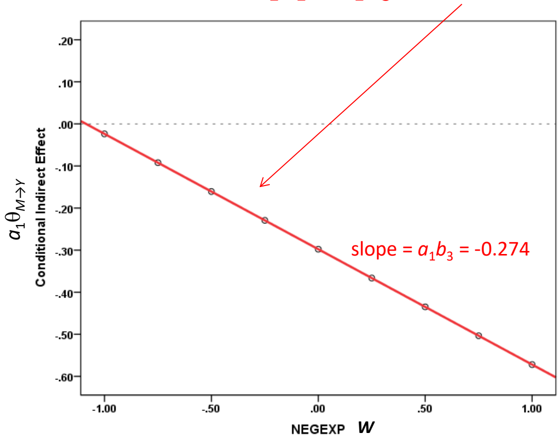
$a_1\theta_{M\to Y} = a_1(b_1 + b_3W)$
$= 0.610(-0.489 - 0.450W)$
$= a_1b_1 + a_1b_3W$
$= -0.298 - 0.274W$

PROCESS sees that the moderator is continuous so, without instruction otherwise, prints the conditional indirect effect at the 16th, 50th, and 84th percentiles. For the mean +/- 1 SD add `moments = 1` to the command line.

A statistical test of moderated mediation

in the second stage moderated mediation model

$$a_1\theta_{M\to Y} = a_1(b_1 + b_3W) = 0.610(-0.489 - 0.450W)$$
$$= a_1b_1 + a_1b_3W = -0.298 - 0.274W$$



The indirect effect is a function of W (negative nonverbal expressivity) in our model. This function is a line.

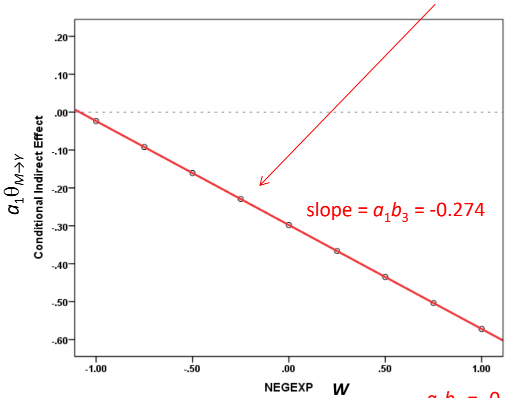
$a_1\theta_{M\to Y} = a_1(b_1 + b_3W)$
$= a_1b_1 + a_1b_3W$
$= -0.298 - 0.274W$

An inference about the slope of this line—the "index of moderated mediation"—is an inference about whether the indirect effect is moderated: Is this slope significantly different from zero? If so, that supports a claim of "moderated mediation."

As a_1b_3 is a product of regression coefficients, an inferential test should respect the nonnormality of the sampling distribution of the product. A bootstrap confidence interval is a good choice.

This test is provided automatically by PROCESS

$a_1\theta_{M\to Y} = a_1(b_1 + b_3W) = 0.610(-0.489 - 0.450W)$
$= a_1b_1 + a_1b_3W = -0.298 - 0.274W$



The indirect effect is a function of V (negative nonverbal expressivity) in our model. This function is a line.

$$
\begin{aligned}
a\theta_{M\to Y} &= a_1(b_1 + b_3W) \\
&= a_1b_1 + a_1b_3W \\
&= -0.298 - 0.274W
\end{aligned}
$$

Output J

a_1b_3 = -0.274, 95% bootstrap CI = -0.683 to -0.024

```
******************* INDEX OF MODERATED MEDIATION ***********************

            Index   SE(Boot)   BootLLCI   BootULCI
Negexp     -.2742     .1727     -.6833     -.0243
```

This slope is statistically different from zero. The indirect effect depends on negative nonverbal expressivity…. The mediation is moderated.

Where is this test discussed?

Multivariate Behavioral Research, 0:1–22, 2015
Copyright © Taylor & Francis Group, LLC
ISSN: 0027-3171 print / 1532-7906 online
DOI: 10.1080/00273171.2014.962683

Routledge
Taylor & Francis Group

An Index and Test of Linear Moderated Mediation

Andrew F. Hayes
The Ohio State University

I describe a test of linear moderated mediation in path analysis based on an interval estimate of the parameter of a function linking the indirect effect to values of a moderator—a parameter that I call the *index of moderated mediation*. This test can be used for models that integrate moderation and mediation in which the relationship between the indirect effect and the moderator is estimated as linear, including many of the models described by Edwards and Lambert (2007) and Preacher, Rucker, and Hayes (2007) as well as extensions of these models to processes involving multiple mediators operating in parallel or in serial. Generalization of the method to latent variable models is straightforward. Three empirical examples describe the computation of the index and the test, and its implementation is illustrated using Mplus and the PROCESS macro for SPSS and SAS.

AN INDEX AND TEST OF LINEAR MODERATED MEDIATION

Empirically substantiating the boundary conditions of one variable's causal effect on another and the mechanism(s) by which that effect operates are recognized as markers of deeper understanding than merely establishing that X affects Y. Successfully answering such questions as "Under what circumstance does X affect Y?" and "How does X affect Y?" adds much merit to one's science and can enhance its impact on a field.

Questions about the contingencies of an effect are often answered statistically through moderation analysis. Assuming continuous Y and dichotomous and/or continuous X and W (the only case considered in this paper), moderation of the effect of X on Y by W is popularly tested by estimating a linear model of the form

$$Y = i_Y + b_1X + b_2W + b_3XW + e_Y$$

where b_1, b_2, and b_3 are estimated regression coefficients, e_Y is an error in estimation, and i_Y is a regression intercept. X's effect on Y is linearly moderated by W if the regression coefficient for XW is different from zero by an inferential test or confidence interval.[1]

Questions about the mechanism by which an effect operates are frequently answered with mediation analysis. In a mediation analysis, focus is on the estimation of the indirect effect of X on Y through an intermediary *mediator* variable M causally located between X and Y (i.e., a model of the form $X \to M \to Y$). Assuming continuous Y and M, the indirect effect of X on Y through M can be derived using two linear models:

$$M = i_M + aX + e_M$$
$$Y = i_Y + c'X + bM + e_Y$$

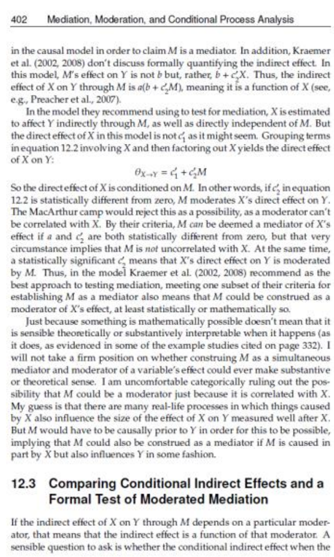
where a, b, and c' are estimated regression coefficients, i_M and i_Y are regression intercepts, and e_M and e_Y are errors in estimation. The product of a and b quantifies the indirect effect of X on Y and estimates how much two cases that differ by one unit on X are estimated to differ on Y through the effect of X on M which in turn influences Y. Evidence that the indirect effect is different from zero by an inferential test or confidence interval bolsters a claim that the effect of X on Y is mediated at least in part by M.

Mediation and moderation analysis can be analytically integrated into a unified statistical model. Although not a new idea by any means—such terms as "mediated moderation" and "moderated mediation" appeared in the literature decades ago (e.g., Baron & Kenny, 1986; James & Brett, 1984; Judd & Kenny, 1981)—it is only recently that a handful of articles in the methodology literature have provided researchers the tools and systematic procedures for answering questions focused on the "when of the how" or the "how of the when." Hayes (2013) introduces the term

Correspondence concerning this article should be addressed to Andrew F. Hayes, Department of Psychology, The Ohio State University, Columbus, OH 43210, USA. E-mail: hayes.338@osu.edu, Website: www.afhayes.com.

[1] Unless otherwise stated, assume that all regression coefficients and functions thereof in my notation are estimates based on the data available.

Provided with your materials



402 Mediation, Moderation, and Conditional Process Analysis

in the causal model in order to claim M is a mediator. In addition, Kraemer et al. (2002, 2008) don't discuss formally quantifying the indirect effect. In this model, M's effect on Y is not b but, rather, $b + c_2'X$. Thus, the indirect effect of X on Y through M is $a(b + c_2'M)$, meaning it is a function of X (see, e.g., Preacher et al., 2007).

In the model they recommend using to test for mediation, X is estimated to affect Y indirectly through M, as well as directly independent of M. But the direct effect of X in this model is not c_1' as it might seem. Grouping terms in equation 12.2 involving X and then factoring out X yields the direct effect of X on Y:

$$\theta_{X\to Y} = c_1' + c_2'M$$

So the direct effect of X is conditioned on M. In other words, if c_2' in equation 12.2 is statistically different from zero, M moderates X's direct effect on Y. The MacArthur camp would reject this as a possibility, as a moderator can't be correlated with X. By their criteria, M *can* be deemed a mediator of X's effect if a and c_2' are both statistically different from zero, but that very circumstance implies that M is *not* uncorrelated with X. At the same time, a statistically significant c_2' means that X's direct effect on Y is moderated by M. Thus, in the model Kraemer et al. (2002, 2008) recommend as the best approach to testing mediation, meeting one subset of their criteria for establishing M as a mediator also means that M could be construed as a moderator of X's effect, at least statistically or mathematically so.

Just because something is mathematically possible doesn't mean that it is sensible theoretically or substantively interpretable when it happens (as it does, as evidenced in some of the example studies cited on page 332). I will not take a firm position on whether construing M as a simultaneous mediator and moderator of a variable's effect could ever make substantive or theoretical sense. I am uncomfortable categorically ruling out the possibility that M could be a moderator just because it is correlated with X. My guess is that there are many real-life processes in which things caused by X also influence the size of the effect of X on Y measured well after X. But M would have to be causally prior to Y in order for this to be possible, implying that M could also be construed as a mediator if M is caused in part by X but also influences Y in some fashion.

12.3 Comparing Conditional Indirect Effects and a Formal Test of Moderated Mediation

If the indirect effect of X on Y through M depends on a particular moderator, that means that the indirect effect is a function of that moderator. A sensible question to ask is whether the conditional indirect effect when the

Chapter 12 of IMMCPA

PROCESS output

We need an inferential test for these conditional indirect effects. Bootstrap confidence intervals are perfect for the job. PROCESS does this for us automatically.

```
******************** DIRECT AND INDIRECT EFFECTS *************************

Direct effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .3729      .1808     2.0622      .0442      .0100      .7357

INDIRECT EFFECT:
 dysfunc     ->     negtone     ->     perform

     negexp     Effect     BootSE   BootLLCI   BootULCI
     -.5308     -.1523      .1540     -.4335      .1943
     -.0600     -.2813      .1249     -.5432     -.0549
      .6000     -.4623      .1678     -.8095     -.1503

      Index of moderated mediation:
            Index     BootSE   BootLLCI   BootULCI
negexp     -.2742      .1791     -.7172     -.0234
```

Conditional indirect effects with 95% bias-corrected bootstrap CIs based on 10,000 bootstrap samples.

The indirect effect of dysfunctional behavior on performance through negative tone is negative among teams relatively moderate (point estimate: -0.28, 95% CI from -0.54 to -0.05) and relatively high (point estimate: -0.46, 95% CI from -0.81 to -0.14) in negative nonverbal expressivity but not different from zero among those low in negative nonverbal expressivity (point estimate: -0.15, 95% CI from -0.43 to 0.19).

Comparing conditional indirect effects (2nd stage model)

A seemingly sensible question to ask is whether the conditional indirect effect of X when the moderator equals some value $W = w_1$ is different than the conditional indirect effect of X when the moderator is some different value $W = w_2$. For example, is the indirect effect among teams low in negative nonverbal expressivity different from the indirect effect among teams high in negative nonverbal expressivity?

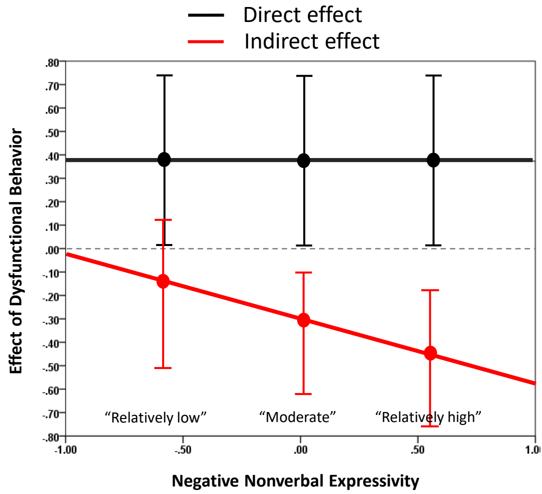
Rejection of the null hypothesis of no moderated mediation based on the index of moderated mediation implies that **any** two conditional indirect effects are different! No additional test is needed.

For example, for the second stage moderated mediation model just estimated:

$$\begin{aligned} a_1(b_1+b_3w_1) - a_1(b_1+b_3w_2) &= a_1(b_1+b_3w_1) - a_1(b_1+b_3w_2) \\ &= a_1b_1+a_1b_3w_1 - a_1b_1 - a_1b_3w_2 \\ &= a_1b_3w_1 - a_1b_3w_2 \\ &= a_1b_3(w_1 - w_2) \end{aligned}$$

If a bootstrap confidence interval for a_1b_3 does not contain zero, then neither will a confidence interval for $a_1b_3(w_1 - w_2)$, ***regardless*** of values of w_1 and w_2 chosen, so long as $w_1 \neq w_2$. And if a bootstrap confidence interval for a_1b_3 contains zero, then so too will a confidence interval for $a_1b_3(w_1 - w_2)$, ***for any two values*** of w_1 and w_2, ($w_1 \neq w_2$).

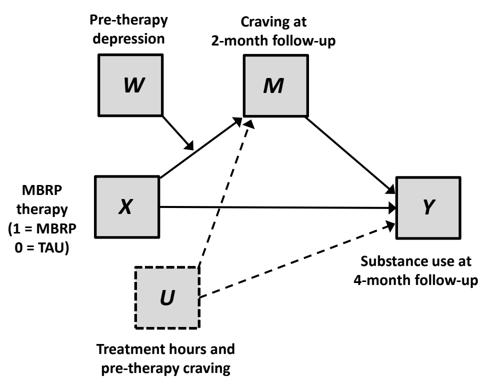
Putting it all together



Point estimates with 95% confidence intervals (OLS confidence intervals for direct effects, bootstrap confidence intervals for indirect effects).

More dysfunctional behavior tends to lead to a more negative work climate, yet this negative climate seems to lower performance only among teams that are more demonstrative of their negative feelings. Such a process does not operate among teams that hide their feelings. Independent of differences between teams in the negative affective tone of the work environment, teams that exhibit more dysfunctional behavior otherwise perform *better*.

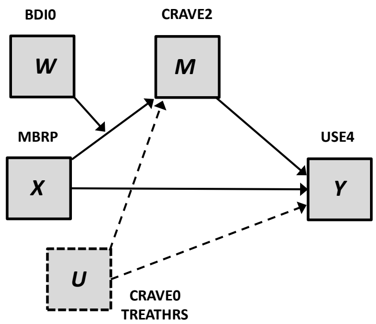
Conditional Process Modeling Example #2



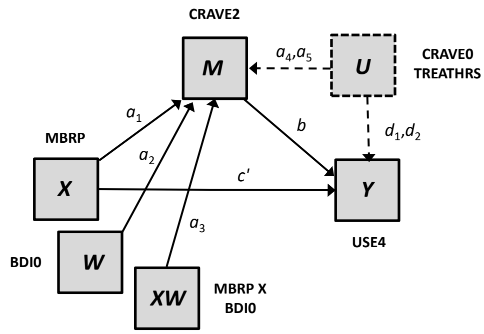
This is a model of **craving (*M*)** as the mechanism by which **mindfulness relapse prevention therapy (*X*)** affects **substance use (*Y*)** relative to therapy as usual. In this model, moderation of the mechanism is proposed as operating in the "first stage" of the mediation process via the moderation of the effect of mindfulness relapse prevention therapy on craving by **pre-therapy depression level (*W*)**.

Conceptual and Statistical Models

Conceptual model



Statistical model



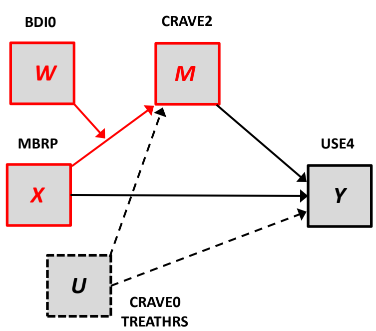
$$\hat{M} = i_1 + a_1X + a_2W + a_3XW + a_4U_1 + a_5U_2$$

$$\hat{Y} = i_2 + c'X + bM + d_1U_1 + d_2U_2$$

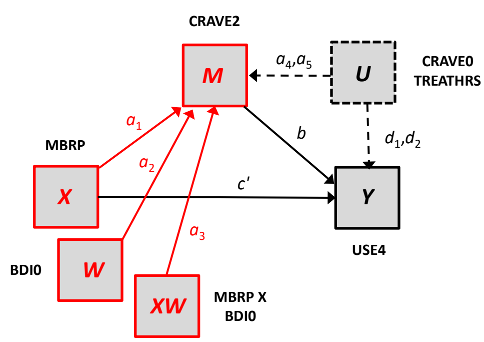
The number of equations needed is equal to the number of variables with arrows pointing at them in the conceptual or statistical diagram.

Conceptual and Statistical Models

Conceptual model



Statistical model



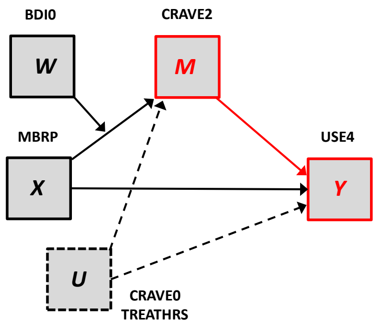
$$\hat{M} = i_1 + \boxed{a_1X + a_2W + a_3XW} + a_4U_1 + a_5U_2$$

$$\hat{Y} = i_2 + c'X + bM + d_1U_1 + d_2U_2$$

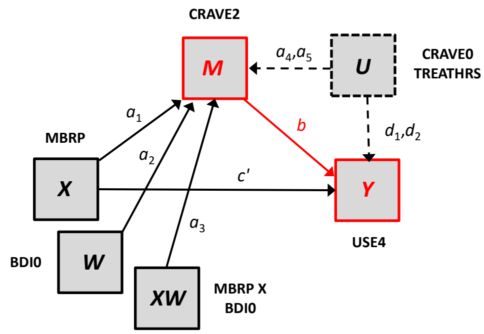
The moderation of the effect of mindfulness behavioral relapse prevention therapy relative to therapy as usual on craving by pre-therapy depression level.

Conceptual and Statistical Models

Conceptual model



Statistical model



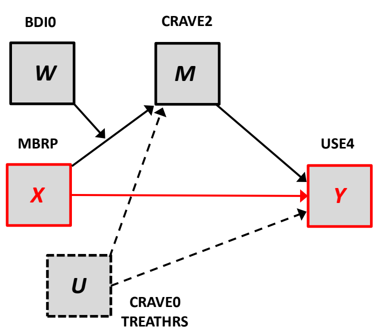
$$\hat{M} = i_1 + a_1X + a_2W + a_3XW + a_4U_1 + a_5U_2$$

$$\hat{Y} = i_2 + c'X \boxed{+bM} + d_1U_1 + d_2U_2$$

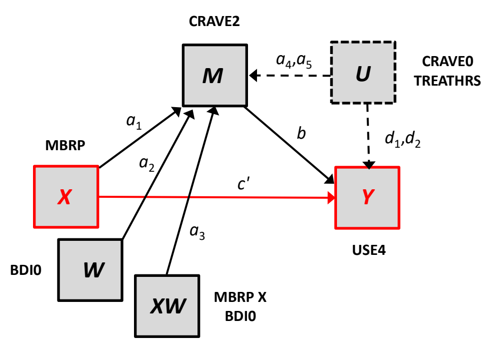
The effect of craving at two month follow-up on substance use after 4 months.

Conceptual and Statistical Models

Conceptual model



Statistical model



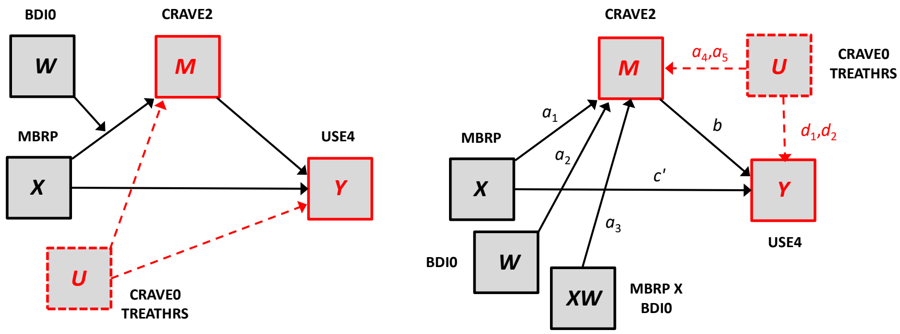
$$\hat{M} = i_1 + a_1X + a_2W + a_3XW + a_4U_1 + a_5U_2$$

$$\hat{Y} = i_2 \boxed{+c'X} + bM + d_1U_1 + d_2U_2$$

The direct effect of mindfulness behavioral relapse prevention therapy on substance use at four month follow up accounting for the mechanism through craving.

Conceptual and Statistical Models

Conceptual model **Statistical model**



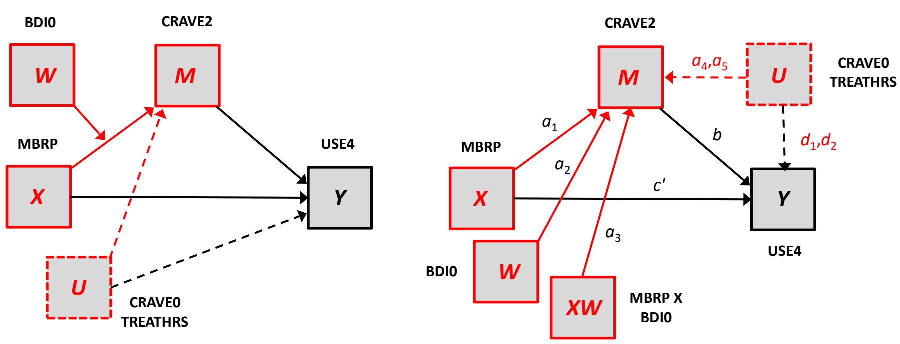
$$\hat{M} = i_1 + a_1X + a_2W + a_3XW + \boxed{a_4U_1 + a_5U_2}$$
$$\hat{Y} = i_2 + c'X + bM + \boxed{d_1U_1 + d_2U_2}$$

Covariates to account for potential confounding by treatment hours and pre-therapy craving.

Estimating the moderation component of the model

The conceptual model proposes that the effect of mindfulness behavioral relapse prevention therapy depends on pre-therapy depression.

Conceptual model **Statistical model**



$$\hat{M} = i_1 + a_1X + a_2W + a_3XW + a_4U_1 + a_5U_2$$

We most care about the moderation components of the model of *M*: a_1 and a_3

We did this already

```
compute mbrpdep = mbrp*bdi0.
regression/dep = crave2/method = enter mbrp bdi0 mbrpdep treathrs crave0.
```

```
data mbrp;set mbrp;mbrpdep=mbrp*bdi0;run;
proc reg data=mbrp;model crave2=mbrp bdi0 mbrpdep treathrs crave0;run;
```

X **= MBRP**
W **= BDI0**
Y **= CRAVE2**

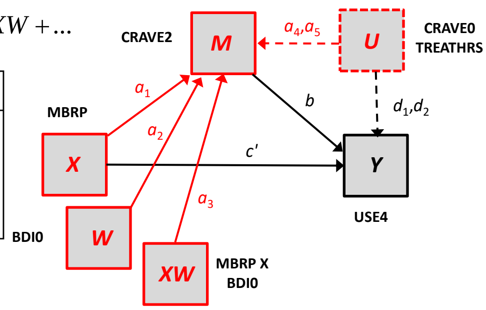
$$\hat{Y} = 1.038 + 0.587X + 1.122W - 0.948XW + \ldots$$

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.038	.470		2.209	.029
	MBRP: Therapy as usual (0) or MBRP therapy (1)	.587	.524	.299	1.120	.264
	BDI0: Beck Depression Inventory baseline	1.122	.276	.366	4.063	.000
	mbrpdep	-.948	.423	-.598	-2.240	.026
	TREATHRS: Hours of therapy	-.018	.010	-.120	-1.719	.088
	CRAVE0: Baseline craving	.192	.073	.183	2.614	.010

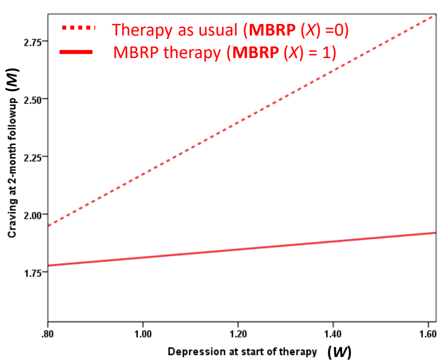
a. Dependent Variable: CRAVE2: Craving at two month follow-up

$a_1 = 0.587$, $a_3 = -0.948$



Pre-therapy depression moderates the effect of mindfulness behavioral relapse prevention therapy on craving. We can say that this moderation/interaction is statistically significant, but this doesn't matter for our purposes because neither the direct nor indirect effects in this model are determined entirely by a_3, and it is the direct and indirect effects we care about. We need a_1 and a_3 to estimate the indirect effects.

Recall the pattern from the earlier analysis



The conditional effect of MBRP therapy ($\theta_{X \to M}$) is defined by the function

$$\boldsymbol{\theta_{X \to M} = 0.587 - 0.948W}$$

BDI0 (W)	$\theta_{X \to M}$
0.877	-0.245
1.196	-0.547
1.515	-0.850

Recall these from our Implementation of the pick-a-point approach.

$$\hat{M} = 1.038 + 0.587X + 1.122W - 0.948XW + \ldots$$

which can be written as

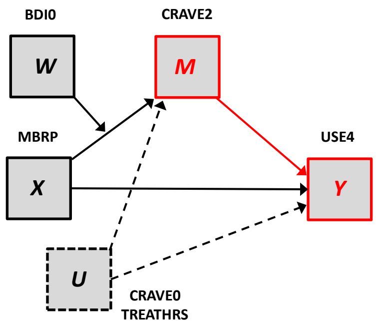
$$\hat{M} = 1.038 + (0.587 - 0.948W)X + 1.122W + \ldots$$

or

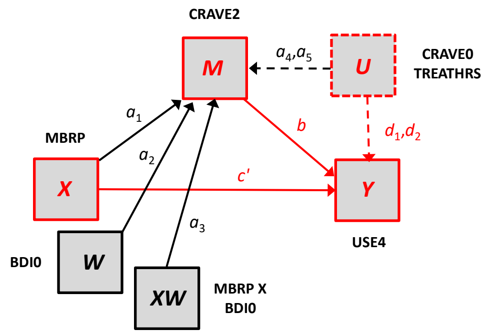
$$\hat{M} = 1.038 + \theta_{X \to M} X + 1.122W + \ldots \text{ where } \quad \theta_{X \to M} = 0.587 - 0.948W = a_1 + a_3 W$$

Estimating the b and c' paths

Conceptual model



Statistical model



$$\hat{M} = i_1 + a_1X + a_2W + a_3XW + a_4U_1 + a_5U_2$$

$$\hat{Y} = i_2 + \boxed{c'X + bM} + d_4U_1 + d_5U_2$$

Estimating the b and c' paths

```
regression/dep=use4/method=enter crave2 mbrp crave0 treathrs.
```

```
proc reg data=mbrp;model use4 = crave2 mbrp crave0 treathrs;run;
```

Coefficients[a]

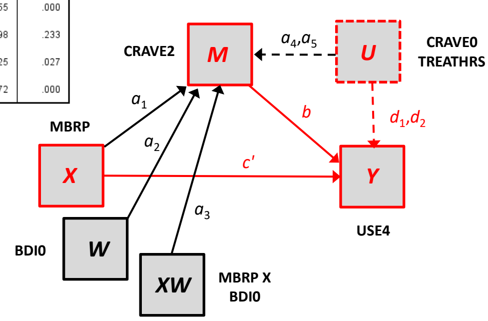
Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.130	.215		5.254	.000
	CRAVE2: Craving at two month follow-up	.481	.040	.710	11.955	.000
	MBRP: Therapy as usual (0) or MBRP therapy (1)	.093	.077	.070	1.198	.233
	CRAVE0: Baseline craving	-.088	.040	-.124	-2.225	.027
	TREATHRS: Hours of therapy	-.020	.006	-.200	-3.572	.000

a. Dependent Variable: USE4: Substance use at four month follow-up

$$\hat{Y} = 1.130 + 0.093X + 0.481M + \ldots.$$

$\boxed{b = 0.481,\ c' = 0.093}$

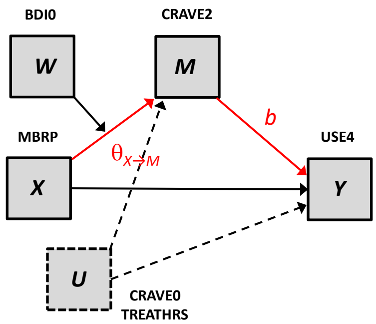
Statistical model



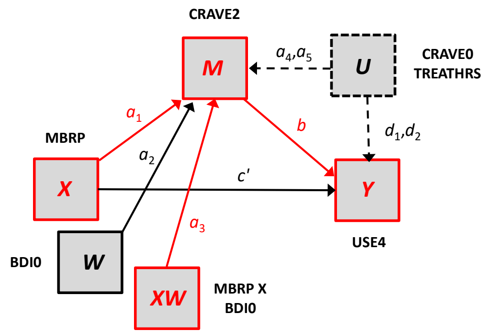
Emphasis in not on statistical significance of the b path, as the indirect effect of X is not defined entirely in terms of b. c' is the direct effect (discussed in a bit).

The conditional indirect effect of *X*

Conceptual model



Statistical model

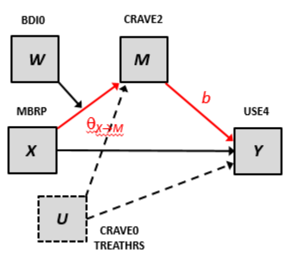


The conditional indirect effect of X on Y through M is the product of the conditional effect of X on M ($\theta_{X\rightarrow M} = a_1 + a_3W$) and effect of M on Y (b):

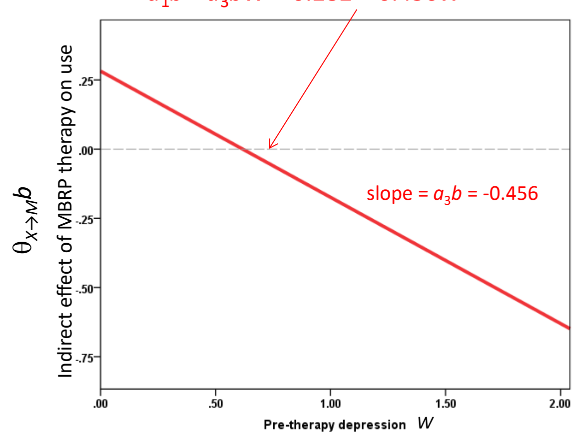
$$\theta_{X\rightarrow M}\, b = (a_1 + a_3W)b = (0.587 - 0.948W)(0.481)$$

The indirect effect of MBRP therapy relative to therapy as usual on later substance use through craving is a function of pre-therapy depression.

A visual representation of the indirect effect



$$\theta_{X\rightarrow M}\, b = (a_1 + a_3W)b = (0.587 - 0.948W)(0.481)$$
$$= a_1b + a_3bW = 0.282 - 0.456W$$

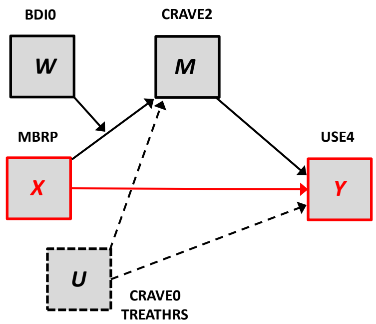


The indirect effect declines with increasing pre-therapy depression. The "**index of moderated mediation**" is a_3b = -0.456. It quantifies the relationship between the moderator and the indirect effect in this model.

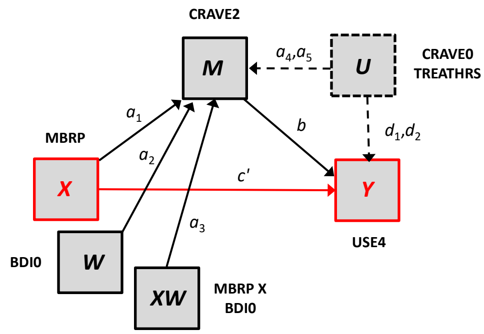
As will be seen, a hypothesis test that the slope of this function is equal to 0 is a formal test of moderated mediation….moderation of the indirect effect.

The direct effect of *X*

Conceptual model



Statistical model



$$\hat{M} = i_1 + a_1X + a_2W + a_3XW + a_4U_1 + a_5U_2$$
$$\hat{Y} = i_2 \boxed{+ c'X +} bM + d_1U_1 + d_2U_2$$

In this model, the direct effect is fixed to be unmoderated. It is a constant rather than a function of another variable in the model. This is a modeling or theoretical decision, not a requirement.

The direct effect of *X* (estimated earlier)

```
regression/dep=use4/method=enter crave2 mbrp crave0 treathrs.
```

```
proc reg data=mbrp;model use4 = crave2 mbrp crave0 treathrs;run;
```

Coefficients[a]

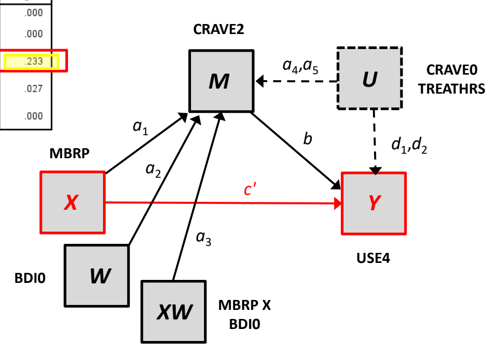
Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.130	.215		5.254	.000
	CRAVE2: Craving at two month follow-up	.481	.040	.710	11.955	.000
	MBRP: Therapy as usual (0) or MBRP therapy (1)	.093	.077	.070	1.198	.233
	CRAVE0: Baseline craving	-.088	.040	-.124	-2.225	.027
	TREATHRS: Hours of therapy	-.020	.006	-.200	-3.572	.000

a. Dependent Variable: USE4: Substance use at four month follow-up

$$\hat{Y} = 1.130 + 0.093X + 0.481M + ...$$

$c' = 0.093$, $t(163) = 1.198$, $p = 0.233$

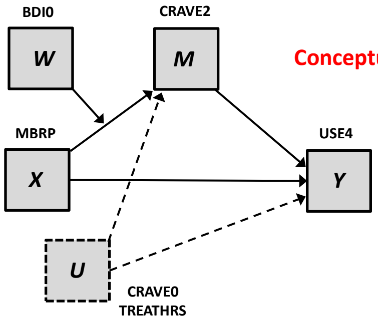
Statistical model



Mindfulness behavioral relapse prevention therapy has no apparent effect on later substance use relative to therapy as usual after accounting for the mechanism through craving.

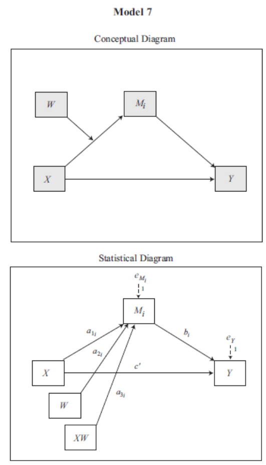
Estimation of the model in PROCESS

PROCESS takes most of the computational burden off our shoulders.



Conceptual model

This is PROCESS model 7

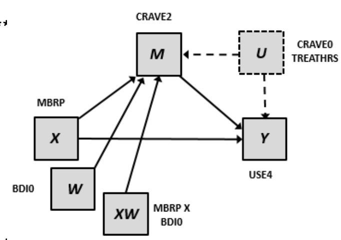



```
process cov = crave0 treathrs/x=mbrp/m=crave2/y=use4/
          w=bdi0/boot=10000/model=7.
```

```
%process (data=mbrp,cov=crave0 treathrs,x=mbrp,m=crave2,y=use4,w=bdi0,
         boot=10000,model=7);
```

PROCESS output

```
**************************************************************************
Model = 7
    Y = use4
    X = mbrp
    M = crave2
    W = bdi0

Statistical Controls:
CONTROL= crave0   treathrs

Sample size
       168

**************************************************************************
Outcome: crave2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5140      .2642      .7277    11.6319     5.0000   162.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.0385      .4701     2.2090      .0286      .1102     1.9668
mbrp          .5872      .5241     1.1204      .2642     -.4478     1.6222
bdi0         1.1221      .2762     4.0625      .0001      .5767     1.6675
int_1        -.9485      .4235    -2.2398      .0265    -1.7847     -.1122
crave0        .1920      .0735     2.6138      .0098      .0470      .3371
treathrs     -.0177      .0103    -1.7190      .0875     -.0380      .0026

Interactions:

 int_1    mbrp        X      bdi0

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



$$\hat{M} = 1.039 + 0.587X + 1.122W - 0.948XW + \ldots$$

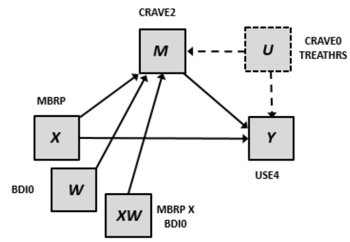
$a_1 = 0.587$

$a_2 = 1.122$

$a_3 = -0.948$

Output K

PROCESS output



$$\hat{Y} = 1.130 + 0.093X + 0.481M + \ldots$$

```
**************************************************************************
Outcome: use4

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .7304      .5335      .2105    46.6070     4.0000   163.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.1298      .2150     5.2545      .0000      .7052     1.5544
crave2        .4810      .0402    11.9547      .0000      .4015      .5604
mbrp          .0926      .0773     1.1979      .2327     -.0601      .2453
crave0       -.0884      .0397    -2.2246      .0275     -.1668     -.0099
treathrs     -.0199      .0056    -3.5720      .0005     -.0309     -.0089
```

$b = 0.481$

$c' = 0.093$

Output K

PROCESS output

```
******************** DIRECT AND INDIRECT EFFECTS *************************

Direct effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .0926      .0773     1.1979      .2327     -.0601      .2453

INDIRECT EFFECT:
 mbrp        ->     crave2       ->     use4

       bdi0     Effect     BootSE   BootLLCI   BootULCI
      .9020     -.1290      .0770     -.2869      .0183
     1.1900     -.2604      .0862     -.4458     -.1090
     1.5180     -.4100      .1367     -.7080     -.1767
```

Direct effect

$c' = .093, p = .232$

W

Output K

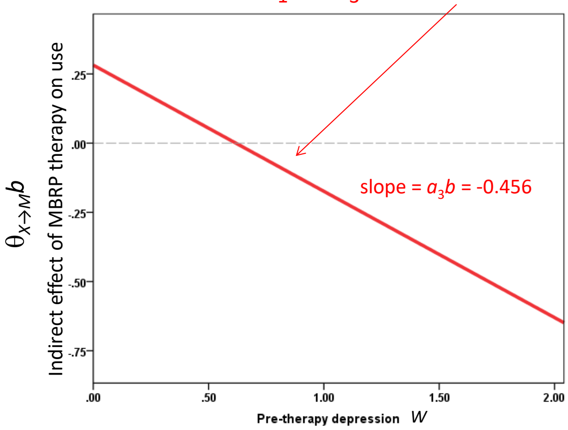
$$\theta_{X \to M}\, b = (a_1 + a_3 W)b = (0.587 - 0.948W)(0.481)$$
$$= a_1 b + a_3 bW = 0.282 - 0.456W$$

PROCESS sees that the moderator is continuous so, without instruction otherwise, prints the conditional indirect effect at 16th, 50th, and 84th percentile of the moderator.

A statistical test of moderated mediation

in the first stage moderated mediation model

$$\theta_{X\to M}\, b = (a_1 + a_3W)b = (0.587 - 0.948W)(0.481)$$
$$= a_1b + a_3bW = 0.282 - 0.456W$$



The indirect effect is a function of W (pre-therapy depression) in our model. This function is a line.

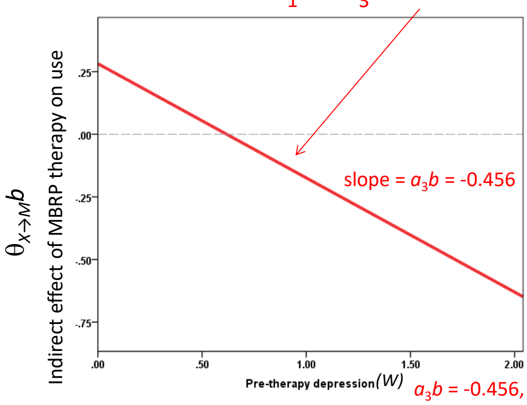
$$\theta_{X\to M}\, b = (a_1 + a_3W)b$$
$$= a_1b + a_3bW$$
$$= 0.282 - 0.456W$$

An inference about the slope of this line—the "index of moderated mediation"—is an inference about whether the indirect effect is moderated: Is this slope significantly different from zero? If so, that supports a claim of "moderated mediation."

As a_3b is a product of regression coefficients, an inferential test should respect the nonnormality of the sampling distribution of the product. A bootstrap confidence interval is a good choice.

This test is provided automatically by PROCESS

$$\theta_{X\to M}\, b = (a_1 + a_3W)b = (0.587 - 0.948W)(0.481)$$
$$= a_1b + a_3bW = 0.282 - 0.456W$$



The indirect effect is a function of W (pre-therapy depression) in our model. This function is a line.

$$\theta_{X\to M}\, b = (a_1 + a_3W)b$$
$$= a_1b + a_3bW$$
$$= 0.282 - 0.456W$$

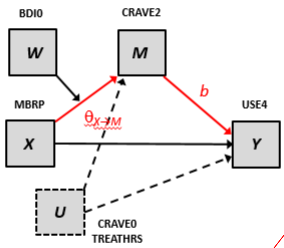
Output K

a_3b = -0.456, 95% bootstrap CI = -0.960 to -0.104

```
******************** INDEX OF MODERATED MEDIATION ************************

Mediator
             Index    SE(Boot)    BootLLCI    BootULCI
crave2      -.4562      .2190      -.9596      -.1044
```

This slope is statistically different from zero. The indirect effect depends on pre-therapy depression....the mediation is moderated.

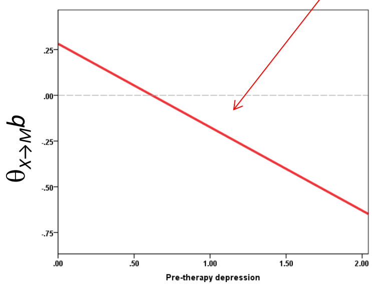
Probing the moderation of mediation



$\theta_{X\to M}\, b = (a_1 + a_3W)b = (0.587 - 0.948W)(0.481)$

The indirect effect decreases with increased pre-therapy depression.

With evidence of moderation of the indirect effect, we can now probe this moderation of mediation through an analogue of the pick-a-point approach used in moderation analysis.



	Pre-therapy depression (W)	$\theta_{X\to M}$	b	$\theta_{X\to M}\, b$
$\bar{W} - SD_W$	0.877	-0.245	0.481	-0.118
$\bar{W}$	1.196	-0.547	0.481	-0.263
$\bar{W} + SD_W$	1.515	-0.850	0.481	-0.409

Conditional indirect effects

We need an inferential test for these conditional indirect effects. Bootstrap confidence intervals are perfect for the job. PROCESS does this for us automatically.

PROCESS output

```
******************** DIRECT AND INDIRECT EFFECTS *************************

Direct effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .0926      .0773     1.1979      .2327     -.0601      .2453

Conditional indirect effect(s) of X on Y at values of the moderator(s):

Mediator
  bdi0     Effect     BootSE   BootLLCI   BootULCI
   .9020    -.1290      .0770     -.2869      .0183
  1.1900    -.2604      .0862     -.4458     -.1090
  1.5180    -.4100      .1367     -.7080     -.1767
```

Conditional indirect effects with 95% bootstrap CIs based on 10,000 bootstrap samples.

Output K

The indirect effect of MBRP therapy relative to therapy as usual on substance use through craving is negative among the relatively moderate (point estimate: -0.260, 95% CI from -0.446 to -0.109) and relatively highly depressed (point estimate: -0.410, 95% CI from -0.708 to -0.177) but not different from zero among the relatively less depressed (point estimate: -0.129, 95% CI from -0.287 to 0.018).

Comparing conditional indirect effects (1st stage model)

A seemingly sensible question to ask is whether the conditional indirect effect of *X* when the moderator equals some value $W = w_1$ is different than the conditional indirect effect of *X* when the moderator is some different value $W = w_2$. For example, is the indirect effect of MBRP therapy through craving different for those relatively low in pre-therapy depression relative those relatively high?

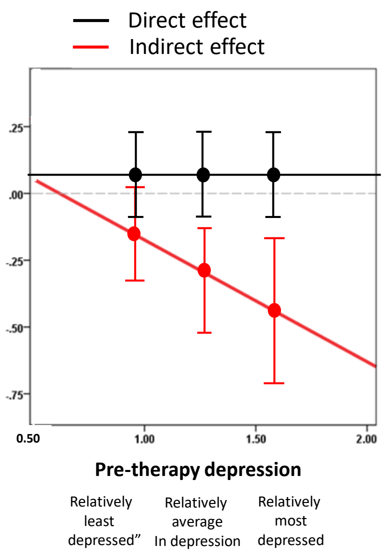
For this model, the difference between any two conditional indirect effects defined by different values of *W* equal to w_1 and w_2 is

$$\begin{aligned}(a_1+a_3w_1)b - (a_1+a_3w_2)b &= a_1b+a_3bw_1 - a_1b - a_3bw_2 \\ &= a_3bw_1 - a_3bw_2 \\ &= a_3b\,(w_1 - w_2)\end{aligned}$$

Rejection of the null hypothesis of no moderated mediation based on the index of moderated mediation implies that **any** two conditional indirect effects are different! No additional test is needed.

If a bootstrap confidence interval for a_3b does not contain zero, then neither will a confidence interval for $a_3b\,(w_1 - w_2)$, ***regardless*** of values of w_1 and w_2 chosen, so long as $w_1 \neq w_2$. And if a bootstrap confidence interval for a_3b contains zero, then so too will a confidence interval for $a_3b\,(w_1 - w_2)$, ***for any two values*** of w_1 and w_2, ($w_1 \neq w_2$).

Putting it all together



MBRP seems to reduce substance use through a reduction in craving which in turn lowers use, but more so among those who are more depressed at the start of therapy. Among those relatively lower in depression, we cannot say definitively that this mechanism is in operation. Independent of this mechanism, there is no evidence of an effect of MBRP therapy on later substance use.

Point estimates with 95% confidence intervals (OLS confidence intervals for direct effects, bootstrap confidence intervals for indirect effects).

A study of sex discrimination in the workplace

Garcia, D. M., Schmitt, M. T., Branscombe, N. R., & Ellemers, N. (2010). Women's reactions to ingroup members who protest discriminatory treatment: The importance of beliefs about inequality and response appropriateness. *European Journal of Social Psychology, 49*, 733-745.

European Journal of Social Psychology
Eur. J. Soc. Psychol. **40**, 733–745 (2010)
Published online 6 July 2009 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/ejsp.644

Research article

Women's reactions to ingroup members who protest discriminatory treatment: The importance of beliefs about inequality and response appropriateness

DONNA M. GARCIA[1]*, MICHAEL T. SCHMITT[2], NYLA R. BRANSCOMBE[3] AND NAOMI ELLEMERS[4]
[1] *University of Guelph, Canada*
[2] *Simon Fraser University, Canada*
[3] *University of Kansas, USA*
[4] *Leiden University, The Netherlands*

Abstract

Our goal was to identify factors that shape women's responses to ingroup members who protest gender discrimination. We predicted and found that women who perceived gender discrimination as pervasive regarded a protest response as being more appropriate than a no protest response and expressed greater liking and less anger towards a female lawyer who protested rather than did not protest an unfair promotion decision. Further, beliefs about the appropriateness of the response to discrimination contributed to evaluations of the protesting lawyer. Perceptions that the complaint was an appropriate response to the promotion decision led to more positive evaluations of an ingroup discrimination protester. Copyright © 2009 John Wiley & Sons, Ltd.

Protest can be an effective means of improving the plight of a devalued group. Historically, there are many examples of protest, even from a single individual, that have advanced a group's social position (e.g. Meritor Savings Bank vs. Vinson, 477 US 57, 1986; Dekker vs. VJV-Centrum ECJ, 1992). Despite the potential gains to be obtained by protesting illegitimate treatment, protestors might not always be appreciated by members of their own group. Whether disadvantaged group members respond positively or negatively to ingroup protestors will likely depend upon the *perceived* implications that the protestor's action has for the ingroup. Unless protest is seen as justified by the social circumstances and an effective means of bringing about positive change, a protestor might be seen as making the ingroup look like complainers. Such threat to the ingroup's reputation could evoke the ire and disdain of the disadvantaged group towards the protestor. Hence, perceptions of the justification for and likely consequences of protest will be critical to others' reactions to an ingroup discrimination claimer. We propose that protest by an ingroup member will be seen as appropriate and thus appreciated to the extent that observers perceive that their ingroup is targeted by pervasive discrimination.

SOCIAL CONSEQUENCES OF CLAIMING DISCRIMINATION

Gender discrimination continues to be widespread throughout Western employment settings (see Charles & Grusky, 2004). The continuation of gender discrimination has substantive negative implications for women's economic and

*Correspondence to: Donna M. Garcia, Department of Psychology, University of Guelph Ontario, Guelph, Ontario, N1G 2W1Canada.
E-mail: donnagarcia3@gmail.com

Received 19 October 2008
Accepted 7 April 2009

All participants (129 females) read a narrative about a female attorney who lost a promotion at her firm to a less qualified male through unequivocally discriminatory actions of the senior partners. The participants where randomly assigned to one of three conditions:

"Individual protest": Participants were told that she protested by describing her qualifications for the job, how much she deserved it, and the unfairness and harm to her own career.

"Collective protest": Participants were told that she protested by describing how the firm has not been fair to women, women are just as qualified as men, and they should be treated equally.

"No protest" : These participants were told that although she was disappointed, she accepted the decision and continued working at the firm.

A study of sex discrimination in the workplace

Garcia, D. M., Schmitt, M. T., Branscombe, N. R., & Ellemers, N. (2010). Women's reactions to ingroup members who protest discriminatory treatment: The importance of beliefs about inequality and response appropriateness. *European Journal of Social Psychology, 49*, 733-745.

European Journal of Social Psychology
Eur. J. Soc. Psychol. **40**, 733–745 (2010)
Published online 6 July 2009 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/ejsp.644

Research article

Women's reactions to ingroup members who protest discriminatory treatment: The importance of beliefs about inequality and response appropriateness

DONNA M. GARCIA[1]*, MICHAEL T. SCHMITT[2], NYLA R. BRANSCOMBE[3] AND NAOMI ELLEMERS[4]
[1] *University of Guelph, Canada*
[2] *Simon Fraser University, Canada*
[3] *University of Kansas, USA*
[4] *Leiden University, The Netherlands*

Abstract

Our goal was to identify factors that shape women's responses to ingroup members who protest gender discrimination. We predicted and found that women who perceived gender discrimination as pervasive regarded a protest response as being more appropriate than a no protest response and expressed greater liking and less anger towards a female lawyer who protested rather than did not protest an unfair promotion decision. Further, beliefs about the appropriateness of the response to discrimination contributed to evaluations of the protesting lawyer. Perceptions that the complaint was an appropriate response to the promotion decision led to more positive evaluations of an ingroup discrimination protester. Copyright © 2009 John Wiley & Sons, Ltd.

Protest can be an effective means of improving the plight of a devalued group. Historically, there are many examples of protest, even from a single individual, that have advanced a group's social position (e.g. Meritor Savings Bank vs. Vinson, 477 US 57, 1986; Dekker vs. VJV-Centrum ECJ, 1992). Despite the potential gains to be obtained by protesting illegitimate treatment, protestors might not always be appreciated by members of their own group. Whether disadvantaged group members respond positively or negatively to ingroup protestors will likely depend upon the *perceived* implications that the protestor's action has for the ingroup. Unless protest is seen as justified by the social circumstances and an effective means of bringing about positive change, a protestor might be seen as making the ingroup look like complainers. Such threat to the ingroup's reputation could evoke the ire and disdain of the disadvantaged group towards the protestor. Hence, perceptions of the justification for and likely consequences of protest will be critical to others' reactions to an ingroup discrimination claimer. We propose that protest by an ingroup member will be seen as appropriate and thus appreciated to the extent that observers perceive that their ingroup is targeted by pervasive discrimination.

SOCIAL CONSEQUENCES OF CLAIMING DISCRIMINATION

Gender discrimination continues to be widespread throughout Western employment settings (see Charles & Grusky, 2004). The continuation of gender discrimination has substantive negative implications for women's economic and

*Correspondence to: Donna M. Garcia, Department of Psychology, University of Guelph Ontario, Guelph, Ontario, N1G 2W1Canada.
E-mail: donnagarcia3@gmail.com

Received 19 October 2008
Accepted 7 April 2009

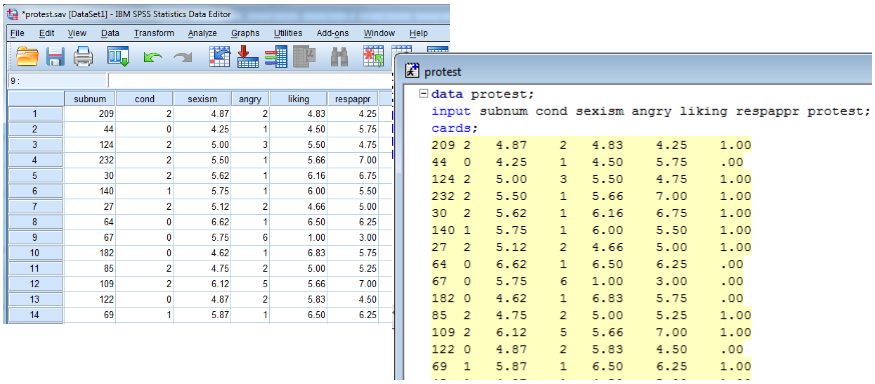
Copyright © 2009 John Wiley & Sons, Ltd.

After reading the narrative, the participants evaluated how **appropriate** they perceived her response to be for the situation. i.e., was it a positive response for dealing with the firm's discrimination.

They also responded to various questions about the attorney that were used to produce a measure of her evaluation which we'll simply call **liking**, such as "I like Catherine," and "Catherine has many positive traits."

Prior to the study, the participants filled out the **modern sexism scale**, used to score each participant with respect to how pervasive she believes sexism and sex discrimination are in society.

The Data: PROTEST



COND: Experimental condition (0 = no protest, 1 = individual protest, 2 = collective protest)

LIKING : Evaluation (liking) of the lawyer (higher = more positive evaluation, i.e. like more)

SEXISM: Score on the modern sexism scale: Beliefs about the pervasiveness of sex discrimination in society (higher = sex discrimination perceived as more pervasive in society)

RESPAPPR: A measure of how appropriate the lawyer's behavior in response to the action of the partners was perceived to be for the situation (higher = more appropriate)

Our objectives with these data

	Perceived Response Appropriateness (*M*)		Liking (*Y*)		
	M	*SD*	Y	*SD*	$\bar{Y}^*$
No protest (*n* = 41)	3.884	1.457	5.310	1.302	5.715
Individual protest (*n* = 43)	5.145	1.075	5.826	0.819	5.711
Collective protest (*n* = 45)	5.494	0.936	5.753	0.936	5.495
All groups combined	4.866	1.348	5.637	1.050	5.637

$\bar{Y}^*$ = adjusted mean, adjusted to the sample mean of perceived response appropriateness.

We'll examine whether the effect of the attorney's response on how she is evaluated (3 conditions)…

… operates through the mechanism of how appropriate her response to the situation is perceived as being (mediation)

Course II….

… depends on perceived pervasiveness of sex discrimination in society (moderation).

… as mediated by the mechanism of perceived appropriateness of her response is moderated by perceived pervasiveness of sex discrimination (moderated mediation)

Single-factor ("one-way") analysis of variance

Did the lawyer's choice as to how to respond (not at all, individual protest, collective protest) influence how she was perceived? That is, is there a difference between conditions, on average, in how much she was liked? Most would answer this using a single-factor (a.k.a. "one-way") analysis of variance (ANOVA).

```
means tables = liking by cond/statistics anova.
```

```
proc anova data=protest;
class cond;model liking = cond;means cond;run;
```

Report

LIKING: liking of the target

COND: experimental condition	Mean	N	Std. Deviation
no protest	5.3102	41	1.30158
individual	5.8260	43	.81943
collective	5.7533	45	.93601
Total	5.6367	129	1.04970

H_0: all means equal

H_a: all means not equal

ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
LIKING: liking of the target * COND: experimental condition	Between Groups	(Combined)	6.523	2	3.262	3.055	.051
	Within Groups		134.515	126	1.068		
	Total		141.039	128			

The lawyer's response to the discrimination affected how much she was liked on average, $F(2,126)$ = 3.055, p = .051. She was most liked when she protested individually (Mean = 5.83, SD = 0.82), next most when protesting collectively (Mean = 5.75, SD = 0.94), and least when she didn't protest at all (Mean = 5.31, SD = 1.30).

Mediation

Does perceived response appropriateness mediate this effect?

	Perceived Response Appropriateness (M)		Liking (Y)		
	M	SD	Y	SD	$\bar{Y}^*$
No protest (n = 41)	3.884	1.457	5.310	1.302	5.715
Individual protest (n = 43)	5.145	1.075	5.826	0.819	5.711
Collective protest (n = 45)	5.494	0.936	5.753	0.936	5.495
All groups combined	4.866	1.348	5.637	1.050	5.637

$\bar{Y}^*$ = adjusted mean, adjusted to the sample mean of perceived response appropriateness.

Here, the presumed cause is a multicategorical variable with three levels. How does one conduct a mediation analysis in such a design?

Mediation of the effect of a multicategorical independent variable

Does perceived response appropriateness mediate this effect?

The causal steps approach

	Perceived Response Appropriateness (*M*)		Liking (*Y*)		
	M	SD	Y	SD	$\bar{Y}^*$
No protest (n = 41)	3.884	1.457	5.310	1.302	5.715
Individual protest (n = 43)	5.145	1.075	5.826	0.819	5.711
Collective protest (n = 45)	5.494	0.936	5.753	0.936	5.495
All groups combined	4.866	1.348	5.637	1.050	5.637

$\bar{Y}^*$ = adjusted mean, adjusted to the sample mean of perceived response appropriateness.

From a single-factor ANOVA,
the effect of experimental condition
on…liking: $F(2,126) = 3.055$, $p = .051$ (**total effect**)
…response appropriateness: $F(2,126) = 22.219$, $p < .001$ (‘**a**’ **effect**)
From a single-factor ANCOVA…the relationship between perceived response appropriateness and liking is positive, $b = 0.412$, $p < 0.001$ (‘***b***’ **effect**), and…
…the effect of condition on liking disappears after controlling for response appropriateness $F(2,125) = 0.729$, $p = .485$ (**direct effect**)

This approach has all the problems of the causal steps (“Baron and Kenny”) approach. There is a better way.

Representing a multicategorical predictor in a linear model

Predictor variables in a linear model must be quantitative or dichotomous (i.e., categorical with only two values). What if we want to include multicategorical variables? **How do we proceed?**

Any multicategorical variable with k categories can be represented with $k - 1$ Variables in a regression model. For example, we can code membership in a category with a set of *dummy variables* and all of these $k - 1$ dummy variables in the model.

Dummy coding (a.k.a. “Indicator coding”)

Set D_1 to 1 for cases in category 1, 0 otherwise
D_2 to 1 for cases in category 2, 0 otherwise
.
.
$D_{(k-1)}$ to 1 for cases in category $k - 1$, 0 otherwise

Category k is called the “reference category,” for reasons that will be clear soon. It is represented here in the coding system, but it doesn’t seem so.

Dummy coding condition

```
frequencies variables = cond.
```

```
proc freq data=protest;tables cond;run;
```

COND: experimental condition

			Frequency	Percent	Valid Percent	Cumulative Percent
cond = 0	Valid	no protest	41	31.8	31.8	31.8
cond = 1		individual	43	33.3	33.3	65.1
cond = 2		collective	45	34.9	34.9	100.0
		Total	129	100.0	100.0	

One possible dummy variable coding system for condition (k = 3 categories):

	D_1	D_2
No protest	0	0
Individual	1	0
Collective	0	1

The "reference category" is the one with zeros on all $k - 1$ dummy variables. In this example, those told the lawyer did not protest are the reference category.

Constructing dummy variables

There is a variety of ways of constructing dummy codes in a computing platform, each with Its dangers, assumptions, and conveniences.

		D_1	D_2
cond = 0	**No protest**	0	0
cond = 1	**Individual**	1	0
cond = 2	**Collective**	0	1

Here is one way

```
compute d1 = 0.
compute d2 = 0.
if (cond = 1) d1 = 1.
if (cond = 2) d2 = 1.
execute.
```

```
data protest;set protest;
d1=0;d2=0;
if (cond=1) then d1 = 1;
if (cond=2) then d2 = 1;
run;
```

This approach can be dangerous. Any cases missing on condition will be coded as if they were assigned to the no protest condition. Use this approach with caution. As a general rule, know your data before you start manipulating it.

Constructing dummy variables

There is a variety of ways of constructing dummy codes in a computing platform, each with its dangers, assumptions, and conveniences.

		D_1	D_2
cond = 0	No protest	0	0
cond = 1	Individual	1	0
cond = 2	Collective	0	1

Here is a safer approach:

```
compute d1 = (cond=1).
compute d2 = (cond=2).
execute.
```

```
data protest;set protest;
d1 = (cond=1);
d2 = (cond=2);
if (cond=.) then d1=.;
if (cond=.) then d2=.;
run;
```

In SPSS, this is very efficient. SAS requires the explicit coding of missing data as such. The SPSS version will leave cases missing on cond missing on d1 and d2.

Estimating liking from experimental condition using regression

```
compute d1 = (cond=1).
compute d2 = (cond=2).
regression/dep = liking/method = enter d1 d2.
```

```
data protest;set protest;
d1 = (cond=1);d2 = (cond=2);run;
proc reg data=protest;model liking = d1 d2;run;
```

We know there are no missing data on **cond**, so this is ok.

$$\hat{Y} = 5.310 + 0.516D_1 + 0.443D_2$$

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.215[a]	.046	.031	1.03324

a. Predictors: (Constant), d2, d1

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.523	2	3.262	3.055	.051[b]
	Residual	134.515	126	1.068		
	Total	141.039	128			

a. Dependent Variable: LIKING: liking of the target

b. Predictors: (Constant), d2, d1

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.310	.161		32.908	.000
	d1	.516	.226	.233	2.287	.024
	d2	.443	.223	.202	1.986	.049

a. Dependent Variable: LIKING: liking of the target

The model reproduces the group means

$$\hat{Y} = 5.310 + 0.516D_1 + 0.443D_2$$

	D_1	D_2
No protest	0	0
Individual	1	0
Collective	0	1

No protest $\hat{Y} = 5.310 + 0.516(0) + 0.443(0) = 5.310 = \overline{Y}_{NP}$

Individual protest $\hat{Y} = 5.310 + 0.516(1) + 0.443(0) = 5.826 = \overline{Y}_{IP}$

Collective protest $\hat{Y} = 5.310 + 0.516(0) + 0.443(1) = 5.753 = \overline{Y}_{CP}$

Report

LIKING: liking of the target

COND: experimental condition	Mean	N	Std. Deviation
no protest	5.3102	41	1.30158
individual	5.8260	43	.81943
collective	5.7533	45	.93601
Total	5.6367	129	1.04970

Interpretation of the coefficients

$$\hat{Y} = i + b_1D_1 + b_2D_2$$
$$\hat{Y} = 5.310 + 0.516D_1 + 0.443D_2$$

	D_1	D_2
No protest	0	0
Individual	1	0
Collective	0	1

i = 5.310 This is the mean liking among those assigned to the no protest condition (D_1 = 0, D_2 = 0).

b_1 = 0.516 This is the mean difference in liking between those in the individual protest condition (D_1 = 1, D_2 = 0) and those in the no protest condition (D_1 = 0, D_2 = 0)

$$b_1 = \overline{Y}_{IP} - \overline{Y}_{NP} = 5.826 - 5.310 = 0.516$$

b_2 = 0.443 This is the mean difference in liking between those in the collective protest condition (D_1 = 0, D_2 = 1) and those in the no protest condition (D_1 = 0, D_2 = 0)

$$b_2 = \overline{Y}_{CP} - \overline{Y}_{NP} = 5.753 - 5.310 = 0.443$$

When D_1 and D_2 are indicator codes constructed in this fashion, b_1 estimates the mean difference in Y between the group coded by D_1 and the reference group, and b_2 estimates the mean difference in Y between the group coded by D_2 and the reference group.

Statistical inference

We are estimating the coefficients of a model of the form

$$\hat{Y} = {}_T i + {}_T b_1 D_1 + {}_T b_2 D_2$$

If there is no actual difference, on average, between these groups on *Y*, this implies that both "true" regression coefficients ${}_T b_1$ and ${}_T b_2$ are both equal to zero. The null hypothesis can be tested by converting the obtained R^2 to an *F*-ratio and then deriving a *p*-value.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.215[a]	.046	.031	1.03324

a. Predictors: (Constant), d2, d1

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.523	2	3.262	3.055	.051[b]
	Residual	134.515	126	1.068		
	Total	141.039	128			

a. Dependent Variable: LIKING: liking of the target

b. Predictors: (Constant), d2, d1

H_0: ${}_T b_1 = {}_T b_2 = 0$

H_a: at least one is different from zero

$$F(k-1, df_{residual}) = \frac{df_{residual} R^2}{(k-1)(1-R^2)}$$

$$F(2,126) = \frac{126(0.046)}{2(1-0.046)}$$

$$F(2,126) = 3.055$$

F(2,126) = 3.055, *p* = .051. Reject H_0. The three group means differ from each other by more than be explained by just 'chance'. Compare this to the one-way ANOVA from earlier.

Statistical inference

We are estimating the coefficients of a model of the form

$$\hat{Y} = {}_T i + {}_T b_1 D_1 + {}_T b_2 D_2$$

b_1 and b_2 can also be used to test hypotheses about differences between groups—specifically, between the group a dummy variable codes and the reference group.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	5.310	.161		32.908	.000
	d1	.516	.226	.233	2.287	.024
	d2	.443	.223	.202	1.986	.049

a. Dependent Variable: LIKING: liking of the target

H_0: ${}_T b_1 = 0$

H_a: ${}_T b_1 \neq 0$

$b_1 = 0.516$, $t(126) = 2.287$, $p = .024$

Those told she individually protested liked her more on average than those told she did not protest.

H_0: ${}_T b_2 = 0$

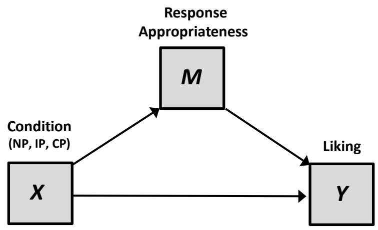
H_a: ${}_T b_2 \neq 0$

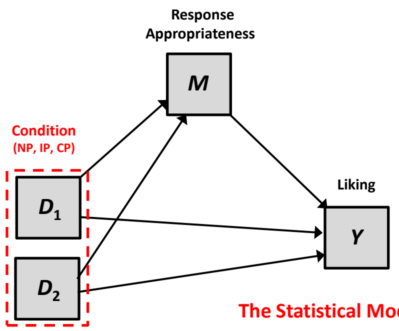
$b_2 = 0.443$, $t(126) = 1.986$, $p = .049$

Those told she collectively protested liked her more on average than those told she did not protest.

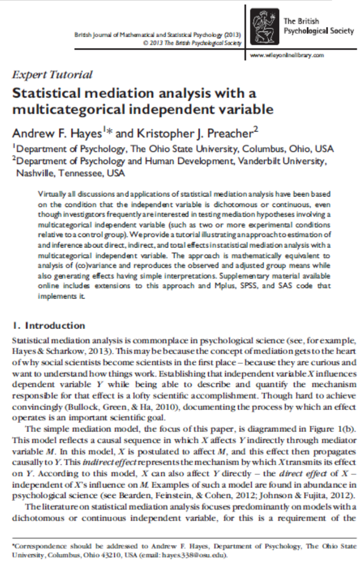
Mediation analysis with a multicategorical independent variable

The Conceptual Model





The Statistical Model


British Journal of Mathematical and Statistical Psychology (2013)
© 2013 The British Psychological Society
The British Psychological Society
www.wileyonlinelibrary.com

Expert Tutorial

Statistical mediation analysis with a multicategorical independent variable

Andrew F. Hayes[1]* and Kristopher J. Preacher[2]
[1]Department of Psychology, The Ohio State University, Columbus, Ohio, USA
[2]Department of Psychology and Human Development, Vanderbilt University, Nashville, Tennessee, USA

Virtually all discussions and applications of statistical mediation analysis have been based on the condition that the independent variable is dichotomous or continuous, even though investigators frequently are interested in testing mediation hypotheses involving a multicategorical independent variable (such as two or more experimental conditions relative to a control group). We provide a tutorial illustrating an approach to estimation of and inference about direct, indirect, and total effects in statistical mediation analysis with a multicategorical independent variable. The approach is mathematically equivalent to analysis of (co)variance and reproduces the observed and adjusted group means while also generating effects having simple interpretations. Supplementary material available online includes extensions to this approach and Mplus, SPSS, and SAS code that implements it.

1. Introduction

Statistical mediation analysis is commonplace in psychological science (see, for example, Hayes & Scharkow, 2013). This may be because the concept of mediation gets to the heart of why social scientists become scientists in the first place – because they are curious and want to understand how things work. Establishing that independent variable *X* influences dependent variable *Y* while being able to describe and quantify the mechanism responsible for that effect is a lofty scientific accomplishment. Though hard to achieve convincingly (Bullock, Green, & Ha, 2010), documenting the process by which an effect operates is an important scientific goal.

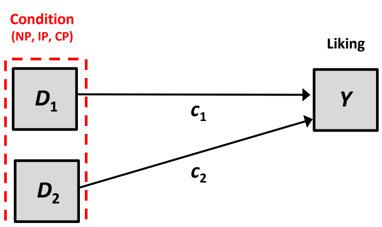
The simple mediation model, the focus of this paper, is diagrammed in Figure 1(b). This model reflects a causal sequence in which *X* affects *Y* indirectly through mediator variable *M*. In this model, *X* is postulated to affect *M*, and this effect then propagates causally to *Y*. This *indirect effect* represents the mechanism by which *X* transmits its effect on *Y*. According to this model, *X* can also affect *Y* directly – the *direct effect* of *X* – independent of *X*'s influence on *M*. Examples of such a model are found in abundance in psychological science (see Bearden, Feinstein, & Cohen, 2012; Johnson & Fujita, 2012).

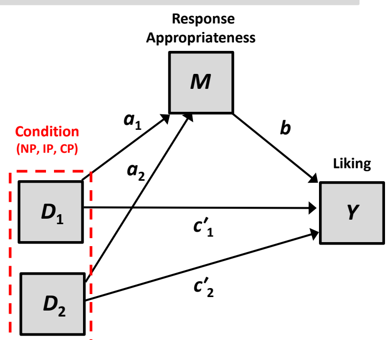
The literature on statistical mediation analysis focuses predominantly on models with a dichotomous or continuous independent variable, for this is a requirement of the

*Correspondence should be addressed to Andrew F. Hayes, Department of Psychology, The Ohio State University, Columbus, Ohio 43210, USA (email: hayes.338@osu.edu).

Hayes and Preacher (2014, BJMSP) available in your materials.

(Relative) total, direct, and indirect effects





c_1 and c_2: *Relative* total effects of experimental condition on liking
c'_1 and c'_2: *Relative* direct effects of experimental condition on liking
a_1b and a_2b: *Relative* indirect effects of condition on liking through perceived response appropriateness.

$$c_1 = c'_1 + a_1b;\ \text{therefore, } a_1b = c_1 - c'_1$$
$$c_2 = c'_2 + a_2b;\ \text{therefore, } a_2b = c_2 - c'_2$$

The relative total effects partition perfectly into relative direct and relative indirect effects. The relative indirect effects are the relative total effects minus the relative direct effects.

Coding the groups

We'll use dummy codes setting the no protest condition to the reference group. Condition (variable name COND) is coded 0 (no protest condition), 1 (individual protest condition), and 2 (collective protest condition).

Condition	D_1	D_2
No protest	0	0
Individual	1	0
Collective	0	1

```
compute d1 = (cond=1).
compute d2 = (cond=2).
execute.
```

```
data protest;set protest;
d1 = (cond=1);
d2 = (cond=2);
if (cond=.) then d1=.;
if (cond=.) then d2=.;
run;
```

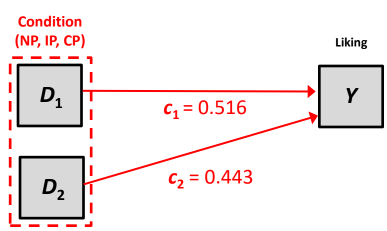
So effects for D_1 will compare individual protest to no protest, and effects for D_2 will compare collective protest to no protest.

The total effect of experimental condition on liking (c paths)

```
regression/dep = liking/method = enter d1 d2.
```

```
proc reg data=protest;model liking=d1 d2;run;
```

We did this already!



Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.215[a]	.046	.031	1.03324

a. Predictors: (Constant), d2, d1

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.523	2	3.262	3.055	.051[b]
	Residual	134.515	126	1.068		
	Total	141.039	128			

a. Dependent Variable: LIKING: liking of the target

b. Predictors: (Constant), d2, d1

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	5.310	.161		32.908	.000
	d1	.516	.226	.233	2.287	.024
	d2	.443	.223	.202	1.986	.049

a. Dependent Variable: LIKING: liking of the target

Relative total effects

Relative to those told she did not protest, those told she individually protested liked her more on average (c_1 = 0.516, p = .024). Relative to those told she did not protest, those told she collectively protested also liked her more on average (c_2 = 0.443, p = .049).

The effect of experimental condition on perceived response appropriateness (*a* paths)

```
regression/dep = respappr/method = enter d1 d2.
```

```
proc reg data=protest;model respappr=d1 d2;run;
```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.511[a]	.261	.249	1.16829

a. Predictors: (Constant), d2, d1

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	60.653	2	30.327	22.219	.000[b]
	Residual	171.977	126	1.365		
	Total	232.631	128			

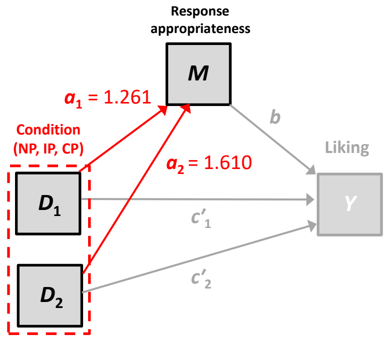
a. Dependent Variable: RESPAPPR: appropriateness of response

b. Predictors: (Constant), d2, d1

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.884	.182		21.288	.000
	d1	1.261	.255	.443	4.946	.000
	d2	1.610	.252	.572	6.384	.000

a. Dependent Variable: RESPAPPR: appropriateness of response



Relative to those told she did not protest, those told she individually protested felt her response was more appropriate on average ($a_1 = 1.261$ $p < .001$). Relative to those told she did not protest, those told she collectively protested felt her response was more appropriate on average ($a_2 = 1.610$, $p < .001$).

The direct effect of condition on liking (*c′* paths)

along with the effect of response appropriateness on liking (*b* path)

```
regression/dep = liking/method = enter respappr d1 d2.
```

```
proc reg data=protest;model liking=respappr d1 d2;run;
```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.503[a]	.253	.235	.91798

a. Predictors: (Constant), d2, RESPAPPR: appropriateness of response, d1

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	35.703	3	11.901	14.123	.000[b]
	Residual	105.336	125	.843		
	Total	141.039	128			

a. Dependent Variable: LIKING: liking of the target

b. Predictors: (Constant), d2, RESPAPPR: appropriateness of response, d1

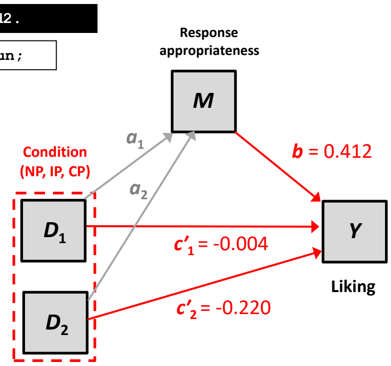
Coefficients[a]

b path

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.710	.307		12.071	.000
	RESPAPPR: appropriateness of response	.412	.070	.529	5.884	.000
	d1	-.004	.219	-.002	-.017	.987
	d2	-.220	.228	-.100	-.966	.336

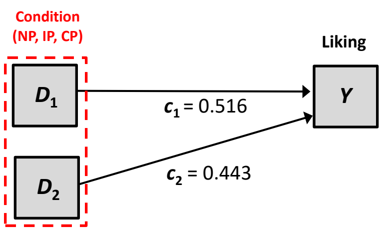
a. Dependent Variable: LIKING: liking of the target

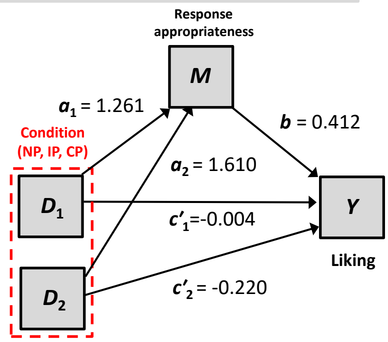
Relative direct effects



Controlling for perceived responses appropriateness, those told she individually protested did not like her any more, on average, than those told she did not protest ($c'_1 = -0.004$, $p = .987$). And those told she collectively protested did not like her any more, on average, than those told she did not protest ($c'_2 = -0.220$, $p = .336$). Holding condition constant, those who perceived her behavior as relatively more appropriate likely her relatively more ($b = 0.412$).

(Relative) total, direct, and indirect effects





c_1 and c_2: *Relative* total effects of condition on liking (c_1 = 0.516, c_2 = 0.443).
c'_1 and c'_2 : *Relative* direct effects of condition on liking (c'_1 = -0.004, c'_2 = -0.220).
a_1b and a_2b: *Relative* indirect effects of condition on liking through perceived response appropriateness
a_1b = 1.261(0.412) = 0.520, a_2b = 1.610(0.412) = 0.663

$c_1 = c'_1 + a_1b$: 0.516 = -0.004 + 1.261(0.412) = -0.004 + 0.520
$c_2 = c'_2 + a_2b$: 0.443 = -0.220 + 1.610(0.412) = -0.220 + 0.663

The relative total effects partition perfectly into relative direct and relative indirect effects. The relative indirect effects are the relative total effects minus the relative direct effects.

Relative total, direct, and indirect effects

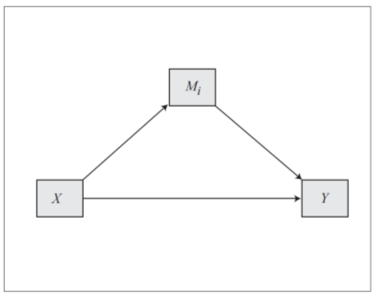
	Perceived Response Appropriateness (M)		Liking (Y)			
	M	SD	$\bar{Y}$	SD	$\bar{Y}^*$	
No protest (n = 41)	3.884	1.457	5.310	1.302	5.715	
a_1 = 1.261		c_1 = 0.516		c'_1=-0.004		
Individual protest (n = 43)	5.145	1.075	5.826	0.819	5.711	c'_2=-0.220
		a_2 = 1.610		c_2 = 0.443		
Collective protest (n = 45)	5.494	0.936	5.753	0.936	5.495	
All groups combined	4.866	1.348	5.637	1.050	5.637	

$\bar{Y}^*$ = adjusted mean, adjusted to the sample mean of perceived response appropriateness.

$c_1 = c'_1 + a_1b$: 0.516 = -0.004 + 1.261(0.412) = -0.004 + 0.520
$c_2 = c'_2 + a_2b$: 0.443 = -0.220 + 1.610(0.412) = -0.220 + 0.663

Estimation using PROCESS

Model 4



PROCESS V3 has an option for specifying X as a multicategorical variable with up to 9 categories. Four options are available for coding the groups.

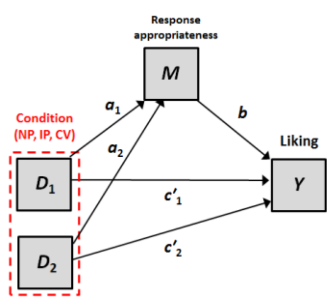
MCX=1 tells PROCESS that X is a multicategorical variable and to use dummy coding to represent the groups. Other coding options are available. See the PROCESS documentation.

MCX	Coding system
1	Simple dummy coding
2	Sequential ("adjacent categories") coding
3	Helmert coding
4	Effect coding

```
process y=liking/m=respappr/x=cond/mcx=1/model=4/total=1/boot=10000.
```

```
%process (data=protest,y=liking,m=respappr,x=cond,mcx=1,model=4,
    total=1,boot=10000);
```

PROCESS output




```
Model  : 4
    Y  : liking
    X  : cond
    M  : respappr

Sample
Size:  129

Coding of categorical X variable for analysis:
   cond     X1     X2
   .000   .000   .000
  1.000  1.000   .000
  2.000   .000  1.000
```

X1 codes individual protest, X2 codes collective protest.
No protest Is the reference group. (The group with the numerically smallest value on the categorical variable is always the reference)

```
**************************************************************************
OUTCOME VARIABLE:
 respappr

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5106      .2607     1.3649    22.2190     2.0000   126.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.8841      .1825    21.2881      .0000     3.5231     4.2452
X1           1.2612      .2550     4.9456      .0000      .7565     1.7659
X2           1.6103      .2522     6.3842      .0000     1.1111     2.1095
```

$$\hat{M} = 3.884 + 1.261 D_1 + 1.610 D_2$$

a_1 path

a_2 path

```
**************************************************************************
OUTCOME VARIABLE:
 liking

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5031      .2531      .8427    14.1225     3.0000   125.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.7103      .3074    12.0711      .0000     3.1020     4.3187
X1           -.0037      .2190     -.0169      .9865     -.4371      .4297
X2           -.2202      .2280     -.9658      .3360     -.6715      .2310
respappr      .4119      .0700     5.8844      .0000      .2734      .5504
```

$$\hat{Y} = 3.710 - 0.004 D_1 - 0.220 D_2 + 0.412 M$$

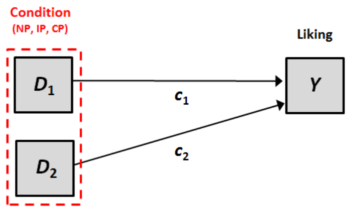
c'_1 path

c'_2 path

b path

Output L

PROCESS output




```
************************* TOTAL EFFECT MODEL ***************************
OUTCOME VARIABLE:
 liking

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .2151      .0463     1.0676     3.0552     2.0000   126.0000      .0506

Model
              coeff         se          t          p       LLCI       ULCI
constant     5.3102      .1614    32.9083      .0000     4.9909     5.6296
X1            .5158      .2255     2.2870      .0239      .0695      .9621
X2            .4431      .2231     1.9863      .0492      .0016      .8845
```

c_1 path

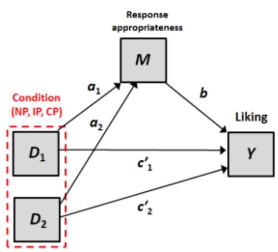
c_2 path

$$\hat{Y} = 5.310 + 0.516D_1 + 0.443D_2$$

Output M

PROCESS output

```
************* TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Relative total effects of X on Y:
       Effect         se          t          p       LLCI       ULCI
X1      .5158      .2255     2.2870      .0239      .0695      .9621
X2      .4431      .2231     1.9863      .0492      .0016      .8845

Omnibus test of total effect of X on Y:
    R2-chng          F        df1        df2          p
      .0463     3.0552     2.0000   126.0000      .0506
----------

Relative direct effects of X on Y
       Effect         se          t          p       LLCI       ULCI
X1     -.0037      .2190     -.0169      .9865     -.4371      .4297
X2     -.2202      .2280     -.9658      .3360     -.6715      .2310

Omnibus test of direct effect of X on Y:
    R2-chng          F        df1        df2          p
      .0087      .7286     2.0000   125.0000      .4846
----------

Relative indirect effects of X on Y

 cond       ->    respappr    ->    liking

       Effect     BootSE   BootLLCI   BootULCI
X1      .5195      .1524      .2590      .8536
X2      .6633      .1671      .3684     1.0187
```



Indirect effect a_1b with bootstrap confidence interval

Indirect effect a_2b with bootstrap confidence interval

Output M

Those told she individually protested liked her more than those told she did not protest because protesting was perceived as more appropriate than not, which in turn enhanced liking (point estimate = 0.520, 95% CI: 0.259 to 0.854). There is no direct effect of individually protesting on liking. Those told she collectively protested liked her more than those told she did not protest because protesting was perceived as more appropriate than not, which in turn enhanced liking (point estimate= 0.663, 95% CI: 0.368 to 1.019). There is no direct effect of collectively protesting on liking.

Omnibus inference

PROCESS gives us tests of the k-1 relative total effects. It also provides a test of equality of the k group means on Y---the "omnibus" total effect. This is equivalent to a single-factor ANOVA.

```
************* TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Relative total effects of X on Y:
        Effect          se           t           p        LLCI        ULCI
X1       .5158       .2255      2.2870       .0239       .0695       .9621
X2       .4431       .2231      1.9863       .0492       .0016       .8845

Omnibus test of total effect of X on Y:
    R2-chng           F         df1         df2           p
      .0463      3.0552      2.0000    126.0000       .0506
----------

Relative direct effects of X on Y
        Effect          se           t           p        LLCI        ULCI
X1      -.0037       .2190      -.0169       .9865      -.4371       .4297
X2      -.2202       .2280      -.9658       .3360      -.6715       .2310

Omnibus test of direct effect of X on Y:
    R2-chng           F         df1         df2           p
      .0087       .7286      2.0000    125.0000       .4846
----------

Relative indirect effects of X on Y

 cond         ->    respappr     ->    liking

        Effect      BootSE    BootLLCI    BootULCI
X1       .5195       .1524       .2590       .8536
X2       .6633       .1671       .3684      1.0187
```

Output M

Test of the "omnibus" total effect.

The three conditions differ on average in liking of the attorney , $F(2,126) = 3.055$, $p = .051$.

Omnibus inference

PROCESS gives us tests of the $k-1$ relative direct effects. It also provides a test of equality of the k group adjusted means on Y when the mediator is held constant---the "omnibus" direct effect. This is equivalent to a single-factor ANCOVA.

Output M

```
************* TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Relative total effects of X on Y:
        Effect          se           t           p        LLCI        ULCI
X1       .5158       .2255      2.2870       .0239       .0695       .9621
X2       .4431       .2231      1.9863       .0492       .0016       .8845

Omnibus test of total effect of X on Y:
    R2-chng           F         df1         df2           p
      .0463      3.0552      2.0000    126.0000       .0506
----------

Relative direct effects of X on Y
        Effect          se           t           p        LLCI        ULCI
X1      -.0037       .2190      -.0169       .9865      -.4371       .4297
X2      -.2202       .2280      -.9658       .3360      -.6715       .2310

Omnibus test of direct effect of X on Y:
    R2-chng           F         df1         df2           p
      .0087       .7286      2.0000    125.0000       .4846
----------

Relative indirect effects of X on Y

 cond         ->    respappr     ->    liking

        Effect      BootSE    BootLLCI    BootULCI
X1       .5195       .1524       .2590       .8536
X2       .6633       .1671       .3684      1.0187
```

Test of the "omnibus" direct effect.

The three conditions do not differ on average in how much they liked her after accounting for group differences in perceived response appropriateness $\Delta R^2 = 0.009$, $F(2,125) = 0.729$, $p = .485$. ΔR^2 is the change in R^2 when the $k-1$ variables coding group are added to the model of Y that already contains the mediator.

Omnibus inference about the indirect effect

- The omnibus tests for the total and direct effect of *X* are not dependent on the system used for coding the groups, even though the relative direct and total effects are.
- The rule that *X* indirectly affects *Y* if at least one relative indirect effect is different from zero means our conclusion will depend on the system used for coding groups, since the relative indirect effects are dependent on that choice.
- If all of the bootstrap confidence intervals for the relative indirect effects straddle zero, that does NOT mean *X* does not indirectly affect *Y*. It could be that a different coding choice produces a different outcome.
- The rule can confirm that *X* indirect affects *Y* if at least one relative indirect effect is different from zero. But a failure to meet this criterion does not disconfirm the existence of an indirect effect of *X* on *Y* through *M*.
- Moral: Choose your coding system wisely, so that it produces relative indirect effects you care about and that are sensitive to the question you are trying to answer.
- There are omnibus tests of the indirect effect that are not sensitive to the coding choice. This must be done in SEM and can require problematic assumptions.

Mediation analysis in the 2-condition within-subject design

Data are from Judd et al. (2001). Estimating and testing mediation and moderation in within-subjects designs. *Psychological Methods, 6*, 115 –134.

20 participants with chronic pain symptoms participated in a pain drug trial. Each was measured twice, once after administration of a placebo **and** once after a administration of a pain inhibiting drug. Order is randomized.

Measurement 1 = Following placebo
Measurement 2 = Following drug

Y = pain sensations (0 to 100, higher = more)
M = pain enhancing hormone (0 to 100, higher = *more*)

Analytical goal: Determine if the effect of the drug on pain experienced operates through the mechanism of reducing pain enhancing hormone levels.

JUDD.sav JUDD.sas

pain1	pain2	hormone1	hormone2
73.00	61.00	37.00	33.00
57.00	55.00	30.00	28.00
57.00	61.00	30.00	36.00
67.00	49.00	31.00	30.00
80.00	75.00	37.00	35.00
56.00	60.00	33.00	34.00
72.00	73.00	38.00	35.00
81.00	65.00	43.00	29.00
61.00	59.00	33.00	31.00
67.00	48.00	20.00	17.00
74.00	64.00	43.00	41.00
70.00	55.00	34.00	27.00
83.00	68.00	41.00	39.00
62.00	61.00	35.00	30.00
49.00	55.00	32.00	32.00
74.00	79.00	35.00	37.00
73.00	60.00	36.00	38.00
46.00	51.00	25.00	24.00
60.00	55.00	26.00	25.00
93.00	71.00	46.00	39.00

Advantages of such a design…

Designs such as this are common:

- Participants might read two different scenarios that vary on some manipulated feature *X* and offer emotional reactions, make predictions about the own behavior, and so forth, in each.
- A therapist might measure certain symptoms and various outcomes when clients arrive in the office for the first time and after a few months of treatment. So *X* is "pre" or "post", i.e., the passage of time.

Some advantages relative to "between-subjects" design and analysis:

- Fewer participants needed. Rather than having *n* people (n_1 in one condition, n_2 in the other), we need about 0.5*n* in total. That saves effort, time, money, labor, and so forth.
- Greater statistical power. Each person serves as his or her own control. "Noise" due to individual differences that increases standard errors in estimates of effects is reduced, sometimes substantially.
- Reduces (but doesn't eliminate) the fundamental problem of causal inference. We don't have to think counterfactually. We know how a person responds in *each* condition rather than having to make an assumption about he or she might have if assigned to the other condition instead.

Judd, Kenny, and McClelland (2001)

Psychological Methods
2001, Vol. 6, No. 2, 115-134

1082-989X/01/$5.00 DOI: 10.1037//1082-989X.6.2.115

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.

The issues of mediation and moderation have received considerable attention in recent years in both basic and applied research (Baron & Kenny, 1986; James & Brett, 1984; Judd & Kenny, 1981b; MacKinnon & Dwyer, 1993). In addition to knowing whether a particular intervention has an effect, the researcher typically wants to know about factors that affect the magnitude of that effect (i.e., moderation) and mechanisms that produce the effect (i.e., mediation). Such knowledge helps in both theory development and intervention application.

To illustrate the difference between mediation and moderation, consider a design in which a researcher is interested in whether students who are taught with a new curriculum (the treatment condition) show higher performance on a subsequent standardized test than students taught under the old curriculum (the control condition). Assuming that a performance difference is found, one might plausibly hypothesize different mediating mechanisms for this effect. The new curriculum might increase students' interest in the subject matter; it might cause students to study harder outside of class; or it might convey the material more clearly. These are alternative reasons why the performance difference is found, that is, alternative mediators of the treatment effect. The researcher might also be interested in factors that affect the magnitude of the difference between performance following the old curriculum and performance following the new one. That difference might be larger or smaller for different types of students or in different types of classrooms or when taught by different kinds of teachers. All of these then are potential moderators of the treatment effect.

It is possible that the same variable may serve as both a mediator and a moderator. For instance, study time might serve both roles. First, as a mediator, the new curriculum might lead to higher performance because it causes students to study more. Second, as a moderator, the treatment might be especially effective for students who spend more time studying.

Procedures for assessing mediation and moderation have been relatively well worked out through ordinary least squares regression and analysis of variance procedures. Mediation is assessed through a four-step procedure (Baron & Kenny, 1986; Judd & Kenny,

Charles M. Judd and Gary H. McClelland, Department of Psychology, University of Colorado at Boulder; David A. Kenny, Department of Psychology, University of Connecticut.

Preparation of this article was partially supported by National Institute of Mental Health Grants R01 MH45049 and R01 MH51964.

Correspondence concerning this article should be addressed to Charles M. Judd, Department of Psychology, University of Colorado, Boulder, Colorado 80309-0345. Electronic mail may be sent to charles.judd@colorado.edu.

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, Kenny, and McClelland (2001)

Data are from Judd et al. (2001). Estimating and testing mediation and moderation in within-subjects designs. *Psychological Methods, 6*, 115 –134.

Analytical goal. Determine if the effect of a pain-inhibiting drug on pain experienced operates through the mechanism of reducing pain-enhancing hormone levels.

(1) On average, is pain following the drug lower than pain following the placebo?

(2) On average, are there fewer pain enhancing hormones in the blood following the drug relative to the placebo?

(3) Does difference in hormone level predict differences in pain experienced?

(4) Do differences in hormone levels account for differences in pain?

Y_1 pain1	Y_2 pain2	M_1 hormone1	M_2 hormone2
73.00	61.00	37.00	33.00
57.00	55.00	30.00	28.00
57.00	61.00	30.00	36.00
67.00	49.00	31.00	30.00
80.00	75.00	37.00	35.00
56.00	60.00	33.00	34.00
72.00	73.00	38.00	35.00
81.00	65.00	43.00	29.00
61.00	59.00	33.00	31.00
67.00	48.00	20.00	17.00
74.00	64.00	43.00	41.00
70.00	55.00	34.00	27.00
83.00	68.00	41.00	39.00
62.00	61.00	35.00	30.00
49.00	55.00	32.00	32.00
74.00	79.00	35.00	37.00
73.00	60.00	36.00	38.00
46.00	51.00	25.00	24.00
60.00	55.00	26.00	25.00
93.00	71.00	46.00	39.00

Application of Judd et al. (2001)

(1) On average, is pain following the drug significantly lower than pain following the placebo?

```
ttest pairs=pain2 pain1.
```

```
proc ttest data=judd;paired pain2*pain1;run;
```

$\overline{Y}_{drug}$ = 61.250, *SD* = 8.583
$\overline{Y}_{placebo}$ = 67.750, *SD* = 11.889

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	pain2	61.2500	20	8.58318	1.91926
	pain1	67.7500	20	11.88929	2.65853

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	pain2 - pain1	-6.50000	9.23665	2.06538	-10.82289	-2.17711	-3.147	19	.005

Pain experienced following the administration of the drug was 6.500 units lower compared to pain experienced following the placebo, $t(19) = -3.147$, $p < .01$.

Application of Judd et al. (2001)

(2) On average, are there fewer pain-enhancing hormones in the blood following the drug relative to the placebo?

```
ttest pairs=hormone2 hormone1.
```

```
proc ttest data=judd;paired hormed2*hormone1;run;
```

$\overline{M}_{drug}$ = 32.000, *SD* = 5.964
$\overline{M}_{placebo}$ = 34.250, *SD* = 6.414

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	hormone2	32.0000	20	5.96481	1.33377
	hormone1	34.2500	20	6.41442	1.43431

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	hormone2 - hormone1	-2.25000	4.10231	.91730	-4.16994	-.33006	-2.453	19	.024

Pain-enhancing hormone levels were 2.25 units lower on average following the administration of the drug compared to following the placebo, $t(19) = -2.453$, $p < .05$.

Application of Judd et al. (2001)

(3) Does difference in hormone level following the drug relative to placebo predict the difference in pain experienced following the drug relative to placebo?

Regress $Y_2 - Y_1$ on both $M_2 - M_1$ and <u>mean centered</u> $(M_2 + M_1)$, as such:

$$Y_2 - Y_1 = i + b_1(M_2 - M_1) + b_2\left((M_2 + M_1) - \frac{\sum_{i=1}^{n}(M_2 + M_1)}{n}\right) + e$$

```
compute ydiff=pain2-pain1.
compute mdiff=hormone2-hormone1.
compute msumc=(hormone2+hormone1)-66.25.
regression/dep=ydiff/method=enter mdiff msumc.
```

```
data judd;set judd;ydiff=pain2-pain1;mdiff=hormone2-hormone1;
msumc=(hormone2+hormone1)-66.25;run;
proc reg data=judd;model ydiff=mdiff msumc;run;
```

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-3.765	2.087		-1.804	.089
	mdiff	1.215	.457	.540	2.658	.017
	msumc	-.065	.160	-.082	-.406	.690

a. Dependent Variable: ydiff

Relatively fewer pain-enhancing hormones following the drug relative to placebo is associated with less pain following the drug relative to placebo, $b_1 = 1.215$, $p < 0.05$.

Why include (M_2 + M_1) in the model?

The impulse is to model $Y_2 - Y_1$ from $M_2 - M_1$ to assess the effect of difference in *M* on difference in *Y*. ***This is appropriate only if the regression weight estimating Y_2 from M_2 is the same as the regression weight estimating Y_1 from M_1.***

If Y_1 and Y_2 are linked to M_1 to M_2 as

$Y_1 = i_1 + dM_1 + e_1$
$Y_2 = i_2 + dM_2 + e_2$, then

$$\begin{aligned} Y_2 - Y_1 &= (i_2 + dM_2 + e_2) - (i_1 + dM_1 + e_1) \\ &= (i_2 - i_1) + d(M_2 - M_1) + (e_2 - e_1) \\ &= i_3 + d(M_2 - M_1) + e_3 \end{aligned}$$

where $i_3 = i_2 - i_1$ and $e_3 = e_2 - e_1$

But if Y_1 and Y_2 are linked to M_1 and M_2 as

$Y_1 = i_1 + d_1M_1 + e_1$
$Y_2 = i_2 + d_2M_2 + e_2$, then

$$\begin{aligned} Y_2 - Y_1 &= (i_2 + d_2M_2 + e_2) - (i_1 + d_1M_1 + e_1) \\ &= (i_2 - i_1) + (d_2M_2 - d_1M_1) + (e_2 - e_1) \end{aligned}$$

It can be shown that this is equivalent to

$$\begin{aligned} Y_2 - Y_1 &= (i_2 - i_1) + 0.5(d_2 - d_1)(M_2 + M_1) + 0.5(d_2 + d_1)(M_2 + M_1) + (e_2 - e_1) \\ &= i_3 + 0.5(d_2 - d_1)(M_2 + M_1) + 0.5(d_2 + d_1)(M_2 - M_1) + e_3 \end{aligned}$$

where $i_3 = i_2 - i_1$ and $e_3 = e_2 - e_1$

$$Y_2 - Y_1 = i + \boxed{b_1(M_2 - M_1)} + b_2\left((M_2 + M_1) - \frac{\sum_{i=1}^{n}(M_2 + M_1)}{n}\right) + e$$

These are nearly the same, with the exception of the centering, which has no effect on b_1. $b_1 = 0.5(d_2 + d_1)$. So b_1 estimates the effect of $M_2 - M_1$ on $Y_2 - Y_1$ without assuming $d_1 = d_2$. It is also the average within period regression weight estimating *Y* from *X*. **$b_1 = d$ if $d_1 = d_2$. Otherwise, b_1 and *d* will likely be different.**

Why mean center M_2 + M_1?

Mean centering ($M_2 + M_1$) yields an intercept that estimates the average difference in *Y* not attributable to differences in *M*.

$$Y_2 - Y_1 = i + b_1(M_2 - M_1) + b_2\left((M_2 + M_1) - \frac{\sum_{i=1}^{n}(M_2 + M_1)}{n}\right) + e$$

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-3.765	2.087		-1.804	.089
	mdiff	1.215	.457	.540	2.658	.017
	msumc	-.065	.160	-.082	-.406	.690

a. Dependent Variable: ydiff

(4) After accounting for differences in hormone levels, there is no statistically significant difference in pain experienced after the drug relative to the placebo, *i* = -3.765, *p* = 0.089.

Observations

(1) This method is squarely rooted in the causal steps tradition to mediation analysis that has been severely 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) on *Y* (pain).
- 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*.

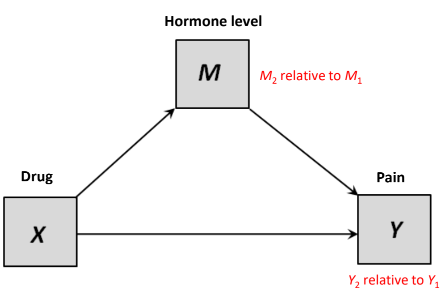
(2) There is no explicit quantification of the indirect effect, but it is the indirect effect that is the primary focus in 21st century mediation analysis.

All of the criticisms of the causal steps approach apply to the Judd, Kenny, and McClelland (2001) method of within-subject mediation analysis.

In a path analytic mediation framework

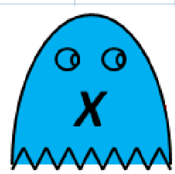
Data are from Judd et al. (2001). Estimating and testing mediation and moderation in within-subjects designs. *Psychological Methods, 6*, 115 –134.

Goal: Model the effect of the pain-facilitating drug on pain sensations, directly as well as indirectly through the effect of the drug on pain-enhancing hormone level.

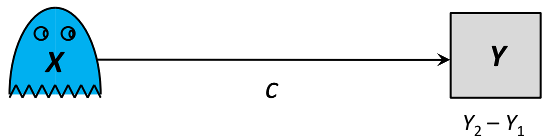


Where is *X* in the data?

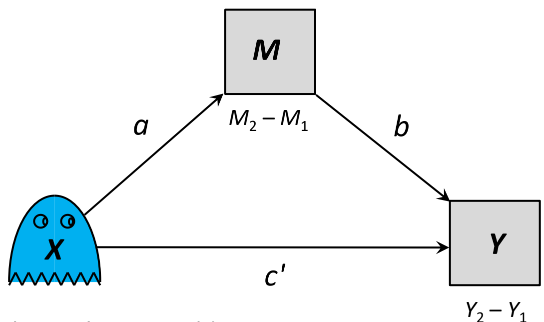
Y_1 pain1	Y_2 pain2	M_1 hormone1	M_2 hormone2
73.00	61.00	37.00	33.00
57.00	55.00	30.00	28.00
57.00	61.00	30.00	36.00
67.00	49.00	31.00	30.00
80.00	[illegible]	[illegible]	35.00
56.00	[illegible]	[illegible]	34.00
72.00	[illegible]	[illegible]	35.00
81.00	[illegible]	[illegible]	29.00
61.00	[illegible]	[illegible]	31.00
67.00	[illegible]	[illegible]	17.00
74.00	[illegible]	[illegible]	41.00
70.00	[illegible]	[illegible]	27.00
83.00	68.00	41.00	39.00
62.00	61.00	35.00	30.00
49.00	55.00	32.00	32.00
74.00	79.00	35.00	37.00
73.00	60.00	36.00	38.00
46.00	51.00	25.00	24.00
60.00	55.00	26.00	25.00
93.00	71.00	46.00	39.00



In a path analytic mediation framework



Y_1 = pain1
Y_2 = pain2
M_1 = hormone1
M_2 = hormone2



$$Y_2 - Y_1 = c + e_1$$

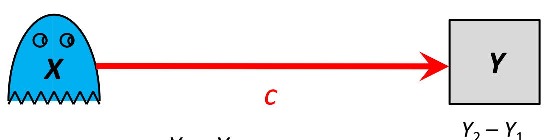
$$M_2 - M_1 = a + e_2$$

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

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

* mean centered

Estimating the total effect (path *c*)



$$Y_2 - Y_1 = c + e_1$$

The total effect is the intercept in a "constant only" model of the difference between pain following the drug (Y_2) and pain following the placebo (Y_1). This is equivalent to the mean difference in pain experienced. No regression analysis needed. Just calculate the mean difference in *Y* (i.e., $Y_2 - Y_1$).

```
compute ydiff=pain2-pain1.
descriptives variables=ydiff.
```

```
data judd;set judd;ydiff=pain2-pain1;run;
proc means data=judd;var ydiff;run;
```

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ydiff	20	-22.00	6.00	-6.5000	9.23665
Valid N (listwise)	20				

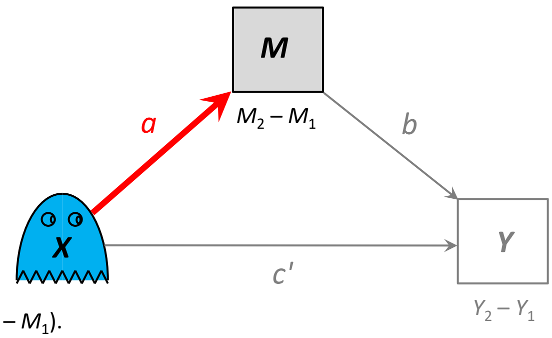
c = -6.500

6.50 units less pain following the drug compared to placebo

Estimating the *a* path

$M_2 - M_1 = a + e_2$

The *a* path is the intercept in a "constant only" model of the difference in hormone following the drug (M_2) and following the placebo (M_1). This is equivalent to the mean difference in hormone level. No regression analysis needed. Just calculate the mean difference in *M* (i.e., $M_2 - M_1$).




```
compute mdiff=hormone2-hormone1.
descriptives variables= mdiff.
```

```
data judd;set judd;mdiff=hormone2-hormone1;run;
proc means data=judd;var mdiff;run;
```

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
mdiff	20	-14.00	6.00	-2.2500	4.10231
Valid N (listwise)	20				

a = -2.250

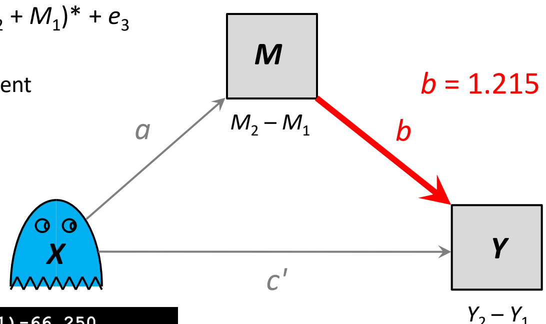
2.50 units less of the pain enhancing hormone following the drug compared to placebo

Estimating the *b* path

* mean centered

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

The *b* path is the regression coefficient for the mean difference in *M* in a model of the mean difference in *Y*, including the mean centered sum of M_1 and M_2 as a covariate. In these data, the mean of the sum of M_1 and M_2 is 66.250.




```
compute msumc=(hormone2+hormone1)-66.250.
regression/dep=ydiff/method=enter mdiff msumc.
```

```
data judd;set judd;mdiff=hormone2-hormone1;msumc=(hormone2+hormone1)-66.250;run;
proc reg data=judd;model ydiff=mdiff msumc;run;
```

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-3.765	2.087		-1.804	.089
	mdiff	1.215	.457	.540	2.658	.017
	msumc	-.065	.160	-.082	-.406	.690

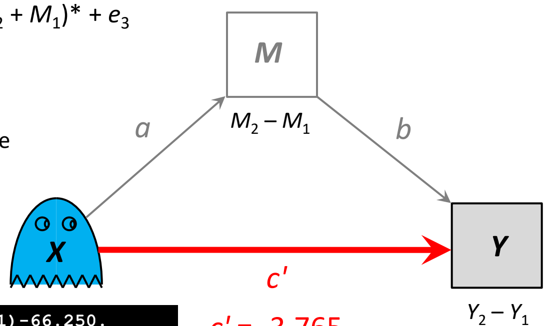
a. Dependent Variable: ydiff

On average, each fewer unit of pain enhancing hormone following drug relative to placebo is associated with 1.215 fewer units of pain following drug relative to placebo.

Estimating the direct effect (path c')

$Y_2 - Y_1 = c' + b(M_2 - M_1) + d_2(M_2 + M_1)^* + e_3$

The direct effect (path c') is the regression constant in the model of the mean difference in Y from the mean difference in M and the mean centered sum of M_1 and M_2. This is the same model used to estimate path b.




```
compute msumc=(hormone2+hormone1)-66.250.
regression/dep=ydiff/method=enter mdiff msumc.
```

$c' = -3.765$

```
data judd;set judd;mdiff=hormone2-hormone1;msumc=(hormone2+hormone1)-66.250;run;
proc reg data=judd;model ydiff=mdiff msumc;run;
```

Coefficients[a]

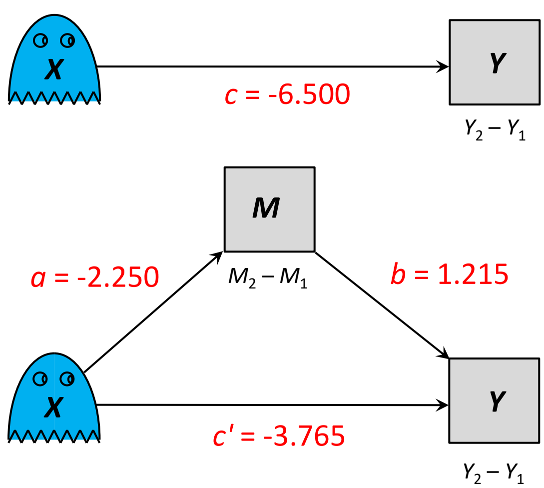
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-3.765	2.087		-1.804	.089
	mdiff	1.215	.457	.540	2.658	.017
	msumc	-.065	.160	-.082	-.406	.690

a. Dependent Variable: ydiff

Independent of the effect of drug vs. placebo hormone difference on pain difference, the drug results in 3.765 units less pain relative to placebo.

Putting it all together

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.



$c = c' + ab$

$$c = \underbrace{-3.765}_{\text{Direct effect}} + \underbrace{(-2.250)(1.215)}_{\text{Indirect effect}} = \underbrace{-3.765}_{\text{Direct effect}} + \underbrace{-2.735}_{\text{Indirect effect}} = \underbrace{-6.500}_{\text{Total effect}}$$

Statistical inference

- **The total effect is just the average difference between Y_2 and Y_1. Inference can proceed in the usual way (paired *t*-test, a confidence interval for the difference, etc.)**

  In these data, c = -6.500, $t(19)$ = -3.147, $p < 0.01$. On average, the drug appears to have an effect on the experience of pain. <u>But we should not insist on evidence of a total effect to proceed with a mediation analysis, as we no longer do elsewhere in modern mediation analysis</u>.

- **The direct effect is just the average difference between Y_2 and Y_1 not accounted for by differences in *M* and their sum. A hypothesis test or confidence interval will do.**

  Here, c' = -3.765, $t(17)$ = -1.804, p = 0.089. Accounting for the effect of change in *M* and individual differences in *M*, the remaining effect of the drug on pain is not statistically different from zero.

- **Whereas the Judd et al. (2001) approach requires the joint significance of *a* and *b* to claim mediation, modern mediation analysis bases claims of mediation on a quantification of the indirect effect and inference about it. The indirect effect is *ab*, which is equivalent to $c - c'$. The sampling distribution of this difference is not normally distributed.**

  ab = -2.250(1.215) = -2.785. This is the reduction in pain due to the drug that results from the effect of the drug on reducing pain-enhancing hormones. A bootstrap CI for inference is a good choice.

MEMORE

This method is described in more detail in Montoya and Hayes (2017*).* The paper includes a description of a new macro (MEMORE, pronounced like “memory”).


Two-Condition Within-Participant Statistical Mediation Analysis:
A Path-Analytic Framework
Amanda K. Montoya and Andrew F. Hayes
The Ohio State University

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

```
memore y=pain2 pain1/m=hormone2 hormone1/samples=10000.
```

```
%memore (data=judd,y=pain2 pain1,m=hormone2 hormone1,samples=10000);
```

MEMORE Output

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

                          Written by Amanda Montoya

                   Documentation available at afhayes.com

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

Variables:
Y = pain2    pain1
M = hormone2 hormone1

Computed Variables:
Ydiff =          pain2      -       pain1
Mdiff =          hormone2   -       hormone1
Mavg =  (        hormone2   +       hormone1 )       /2       Centered

Sample Size:
  20
```

MEMORE constructs differences and averages for you.

```
********************************************************************************
Outcome: Ydiff =  pain2     -       pain1

Model
        Effect         SE          t         df          p       LLCI       ULCI
'X'    -6.5000     2.0654    -3.1471    19.0000      .0053   -10.8233    -2.1767

********************************************************************************
Outcome: Mdiff =  hormone2  -       hormone1

Model
        Effect         SE          t         df          p       LLCI       ULCI
'X'    -2.2500      .9173    -2.4528    19.0000      .0240    -4.1701     -.3299

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

c = -6.500

a = -2.250

MEMORE Output

```
********************************************************************************
Outcome: Ydiff =  pain2     -       pain1

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5554      .3085    65.9384     3.7918     2.0000    17.0000      .0435

Model
           coeff         SE          t         df          p       LLCI       ULCI
'X'      -3.7654     2.0869    -1.8043    17.0000      .0889    -8.1689      .6381
Mdiff     1.2154      .4572     2.6583    17.0000      .0166      .2506     2.1801
Mavg      -.1302      .3209     -.4057    17.0000      .6900     -.8074      .5470

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

Total effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
    -6.5000     2.0654    -3.1471    19.0000      .0053   -10.8233    -2.1767

Direct effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
    -3.7654     2.0869    -1.8043    17.0000      .0889    -8.1689      .6381

Indirect Effect of X on Y through M
          Effect     BootSE   BootLLCI   BootULCI
Ind1     -2.7346     1.3120    -5.6521     -.5230

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′ = -3.765

b = 1.215

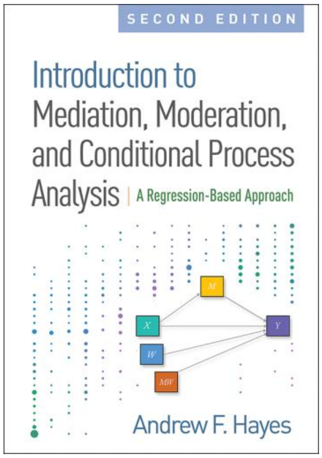
c = -6.500

c′ = -3.765

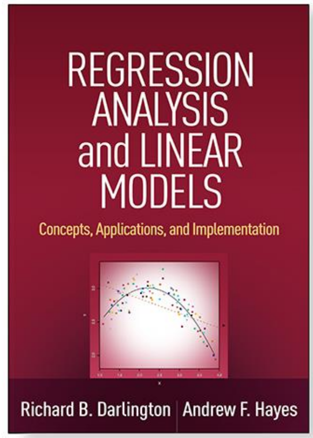
ab with 95% bootstrap confidence interval

Where to learn more

The entire book



Chapters 13, 14, and 15



http://www.guilford.com/

Pertinent Publications

Hayes, A. F., & Rockwood, N. J. (in press). Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour Research and Therapy.*

Hayes, A. F., Montoya, A. K., & Rockwood, N. J. (2017). The analysis of mechanisms and their contingencies: PROCESS versus structural equation modeling. *Australasian Marketing Journal, 25,* 76-81.

Montoya, A. K., & Hayes, A. F. (2017). Two condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods, 22*, 6-27.

Hayes, A. F. (2015). *An index and test of linear moderated mediation*. *Multivariate Behavioral Research, 50,* 1-22.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, *67*, 451-470.

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science, 24*, 1918-1927.

Hayes, A. F., & Preacher, K. J. (2013). Conditional process modeling: Using structural equation modeling to examine contingent causal processes. In G. R. Hancock & R. O. Mueller (Eds.) *Structural equation modeling: A second course* (2nd Ed). Greenwich, CT: Information Age Publishing.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press. (Note: the 2nd edition will be available at the end of 2017)

Pertinent Publications

Hayes, A. F., Glynn, C. J., & Huge, M. E. (2012). Cautions regarding the interpretation of regression coefficients and hypothesis tests in linear models with interactions. *Communication Methods and Measures, 6,* 1-11.

Hayes, A. F., Preacher, K. J., & Myers, T. A. (2011). Mediation and the estimation of indirect effects in political communication research. In E. P. Bucy & R. L. Holbert (Eds), *Sourcebook for political communication research: Methods, measures, and analytical techniques.* (p. 434-465). *New York:* Routledge.

Hayes, A. F., & Preacher, K. J. (2010). Estimating and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate Behavioral Research, 45,* 627-660.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millenium. *Communication Monographs, 76,* 408-420.

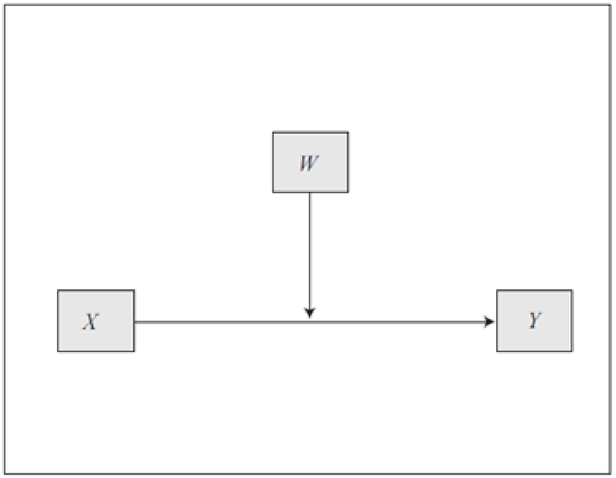
Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods, 41,* 924-936.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879-891.

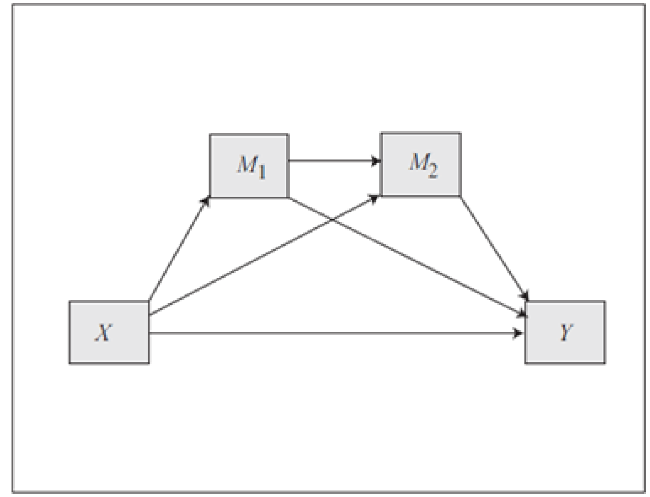
Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Assessing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42,* 185-227.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, and Computers, 36,* 717-731.

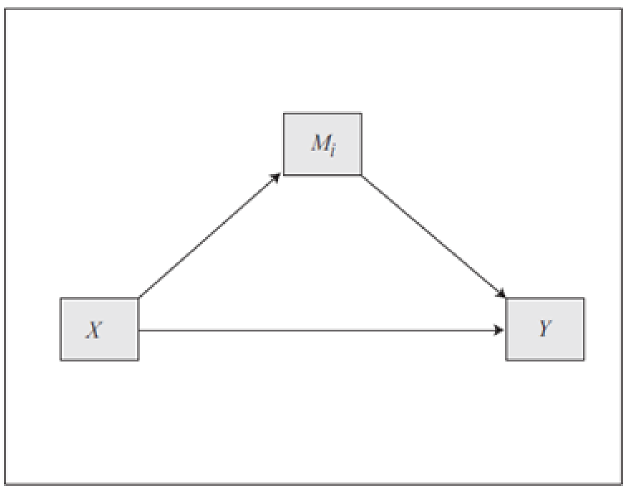
Model 1



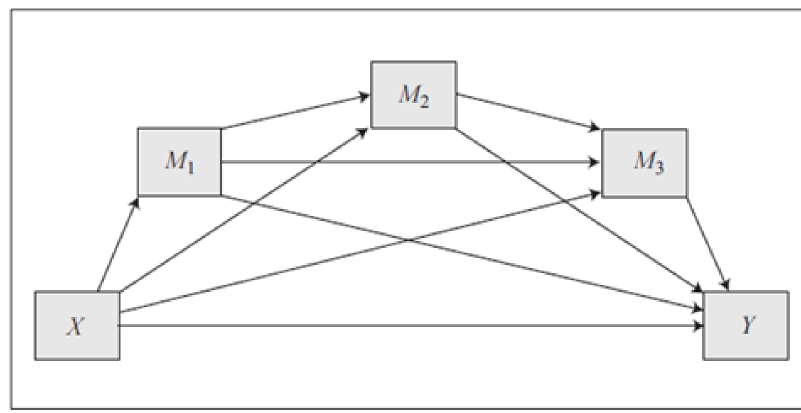
Model 6
(2 mediators)



Model 4

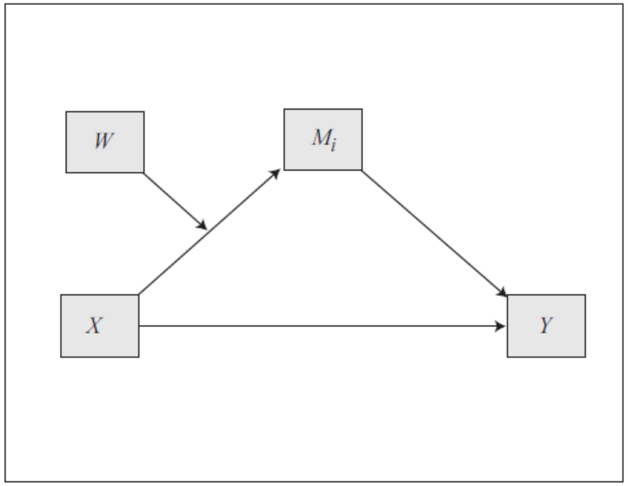


Model 6
(3 mediators)

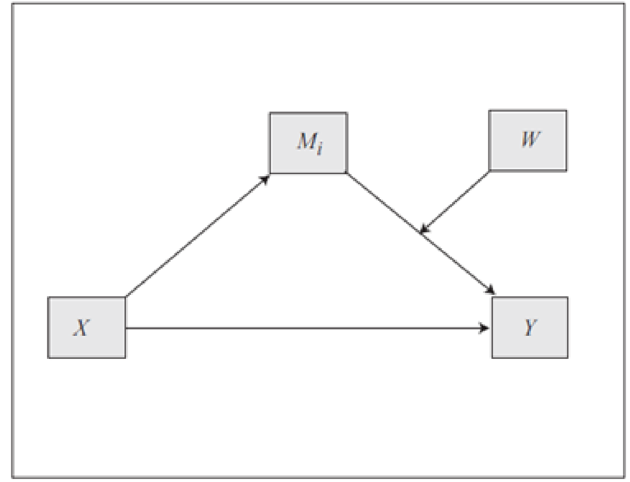


from Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd edition). New York: The Guilford Press   www.guilford.com/p/hayes3 See Appendix A for the complete set of preprogrammed model templates and Appendix B for instructions on how to construct your own models.

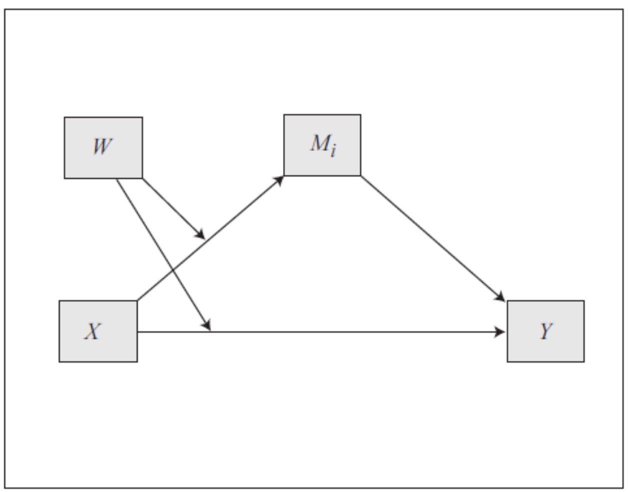
Model 7



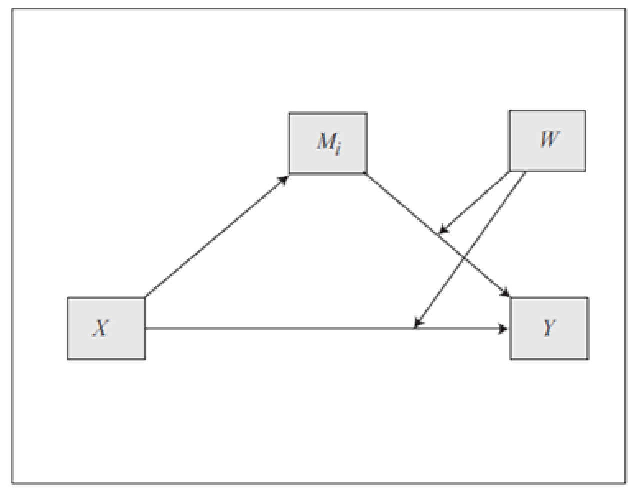
Model 14



Model 8

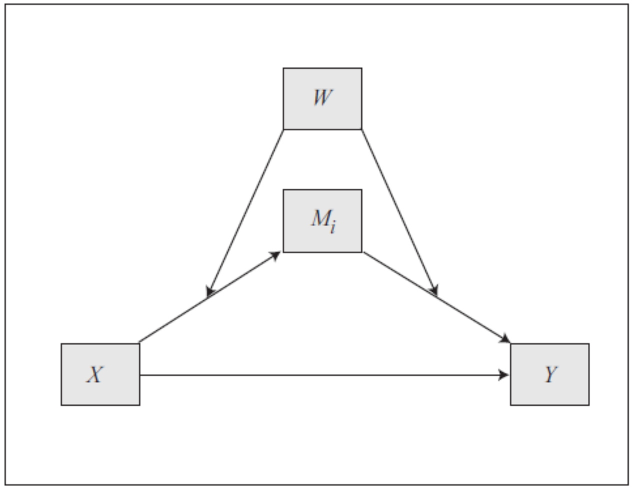


Model 15

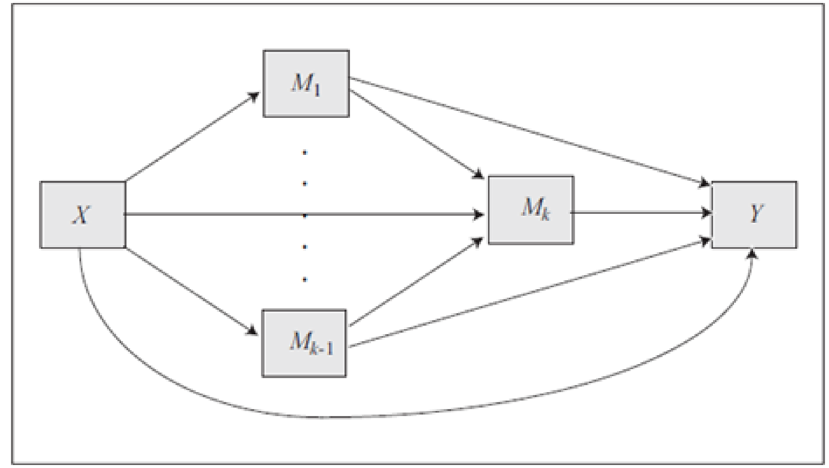


from Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd edition). New York: The Guilford Press www.guilford.com/p/hayes3 See Appendix A for the complete set of preprogrammed model templates and Appendix B for instructions on how to construct your own models.

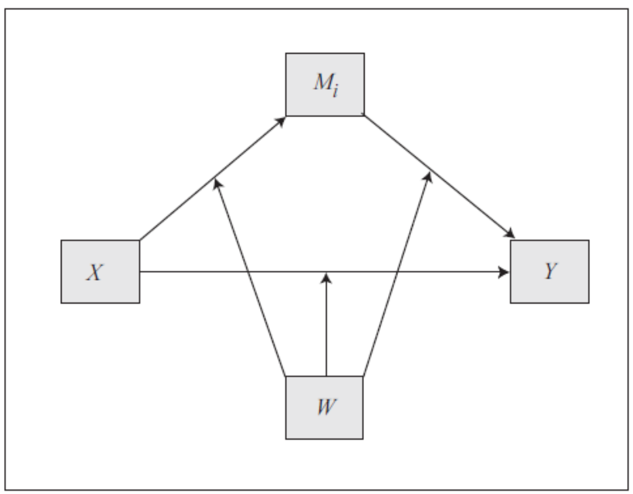
Model 58



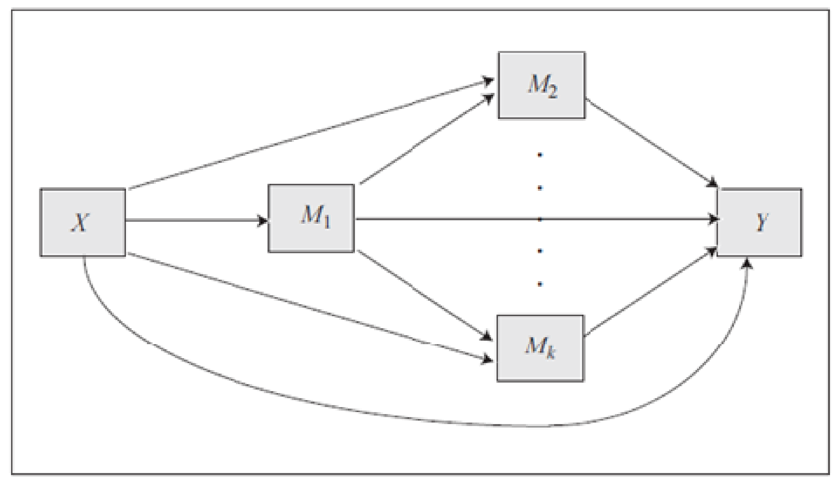
Model 80



Model 59



Model 81



from Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd edition). New York: The Guilford Press   www.guilford.com/p/hayes3 See Appendix A for the complete set of preprogrammed model templates and Appendix B for instructions on how to construct your own models.

OUTPUT A: SAS

The REG Procedure
Model: MODEL1
Dependent Variable: satis1

Number of Observations Read	330
Number of Observations Used	330

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	33.28219	33.28219	87.04	<.0001
Error	328	125.41369	0.38236		
Corrected Total	329	158.69588			

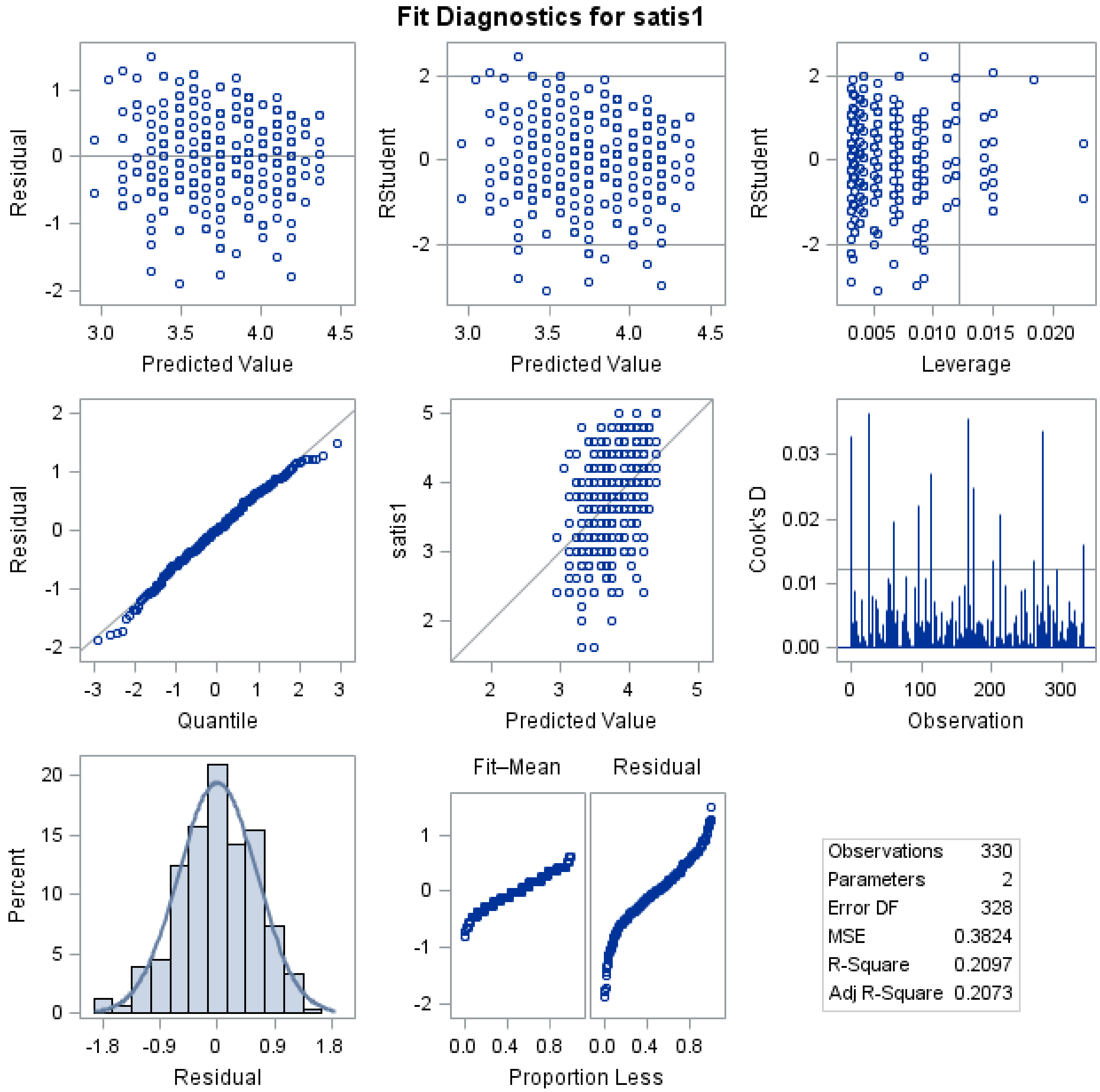
Root MSE	0.61835	**R-Square**	0.2097
Dependent Mean	3.75939	**Adj R-Sq**	0.2073
Coeff Var	16.44817		

Parameter Estimates

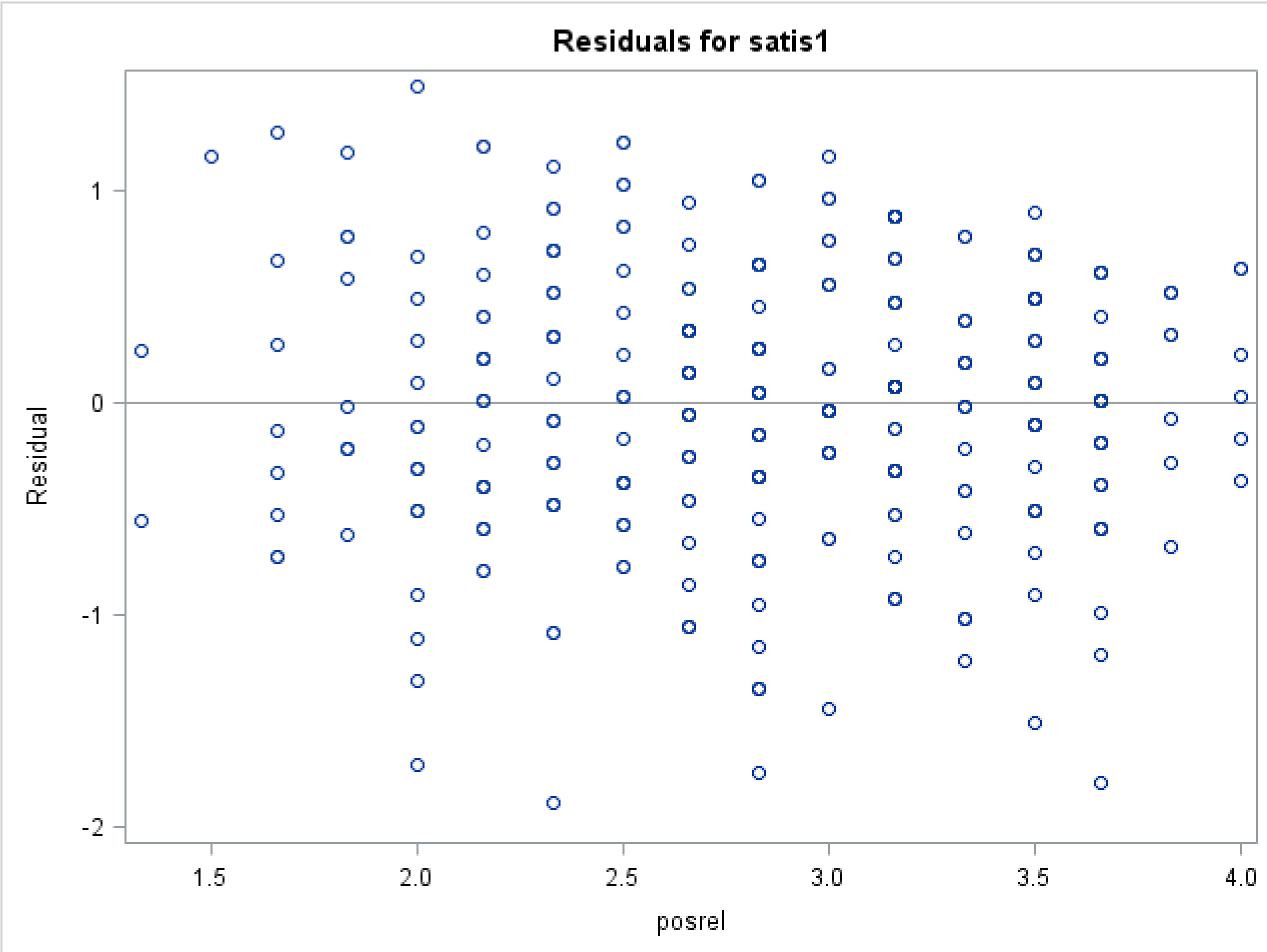
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate	95% Confidence Limits	
Intercept	**1**	2.24622	0.16572	13.55	<.0001	0	1.92021	2.57223
posrel	**1**	0.53076	0.05689	9.33	<.0001	0.45796	0.41884	0.64267

The REG Procedure
Model: MODEL1
Dependent Variable: satis1

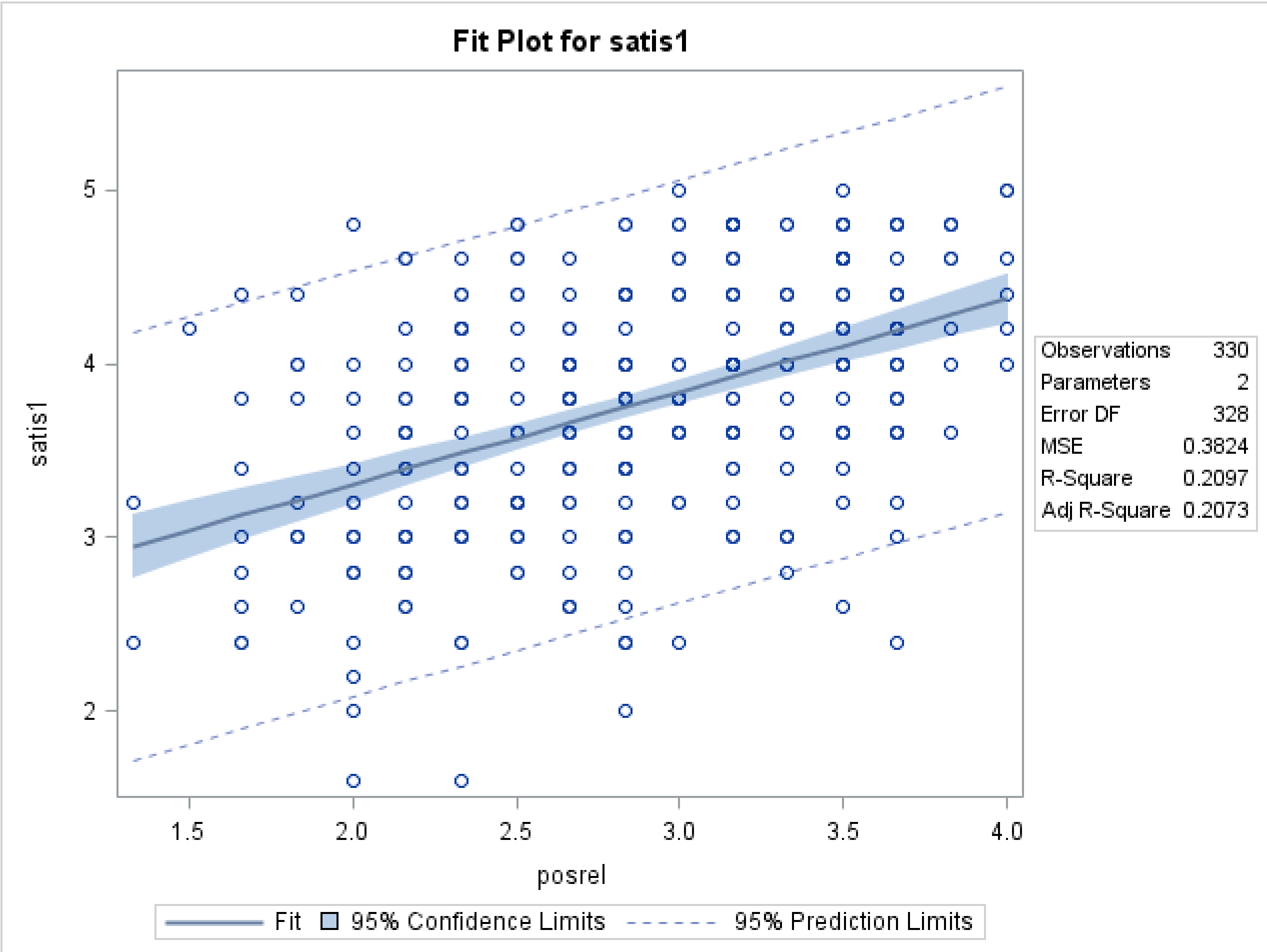
Fit Diagnostics for satis1



Observations	330
Parameters	2
Error DF	328
MSE	0.3824
R-Square	0.2097
Adj R-Square	0.2073



Residuals for satis1
Residual
posrel
1
0
-1
-2
1.5
2.0
2.5
3.0
3.5
4.0



Fit Plot for satis1
satis1
5
4
3
2
1.5
2.0
2.5
3.0
3.5
4.0
posrel
Observations 330
Parameters 2
Error DF 328
MSE 0.3824
R-Square 0.2097
Adj R-Square 0.2073
Fit
95% Confidence Limits
95% Prediction Limits

OUTPUT A: SPSS

```
regression/statistics defaults ci/dep=satis1/method=enter posrel.
```

Regression

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	posrel[b]	.	Enter

a. Dependent Variable: satis1

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.458[a]	.210	.207	.61835

a. Predictors: (Constant), posrel

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	33.282	1	33.282	87.044	.000[b]
	Residual	125.414	328	.382		
	Total	158.696	329			

a. Dependent Variable: satis1

b. Predictors: (Constant), posrel

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.246	.166		13.554	.000
	posrel	.531	.057	.458	9.330	.000

Coefficients[a]

Model		95.0% Confidence Interval for B	
		Lower Bound	Upper Bound
1	(Constant)	1.920	2.572
	posrel	.419	.643

a. Dependent Variable: satis1

OUTPUT B: SAS

The REG Procedure
Model: MODEL1
Dependent Variable: satis1

Number of Observations Read	330
Number of Observations Used	330

Analysis of Variance

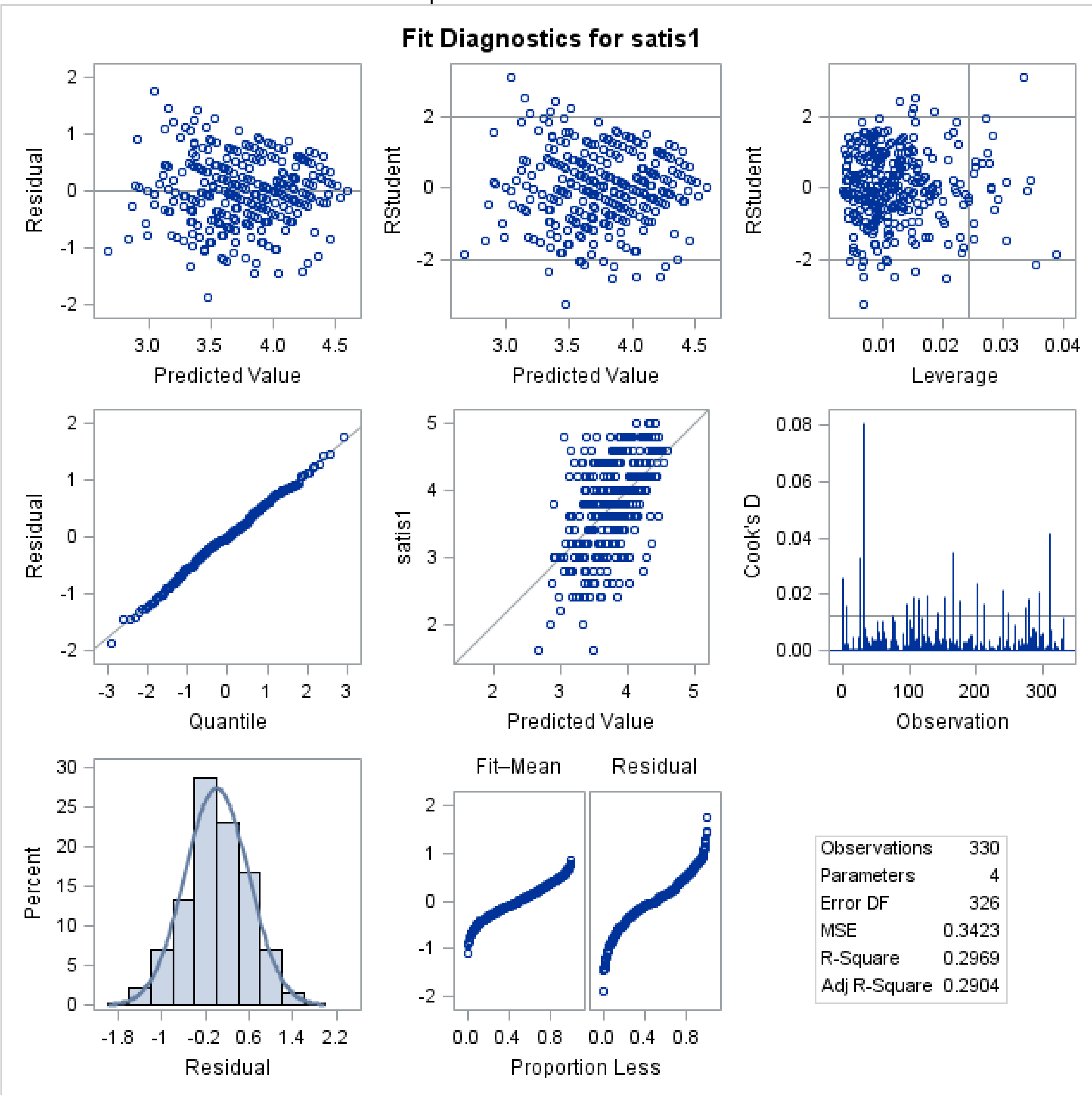
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	47.11953	15.70651	45.89	<.0001
Error	326	111.57635	0.34226		
Corrected Total	329	158.69588			

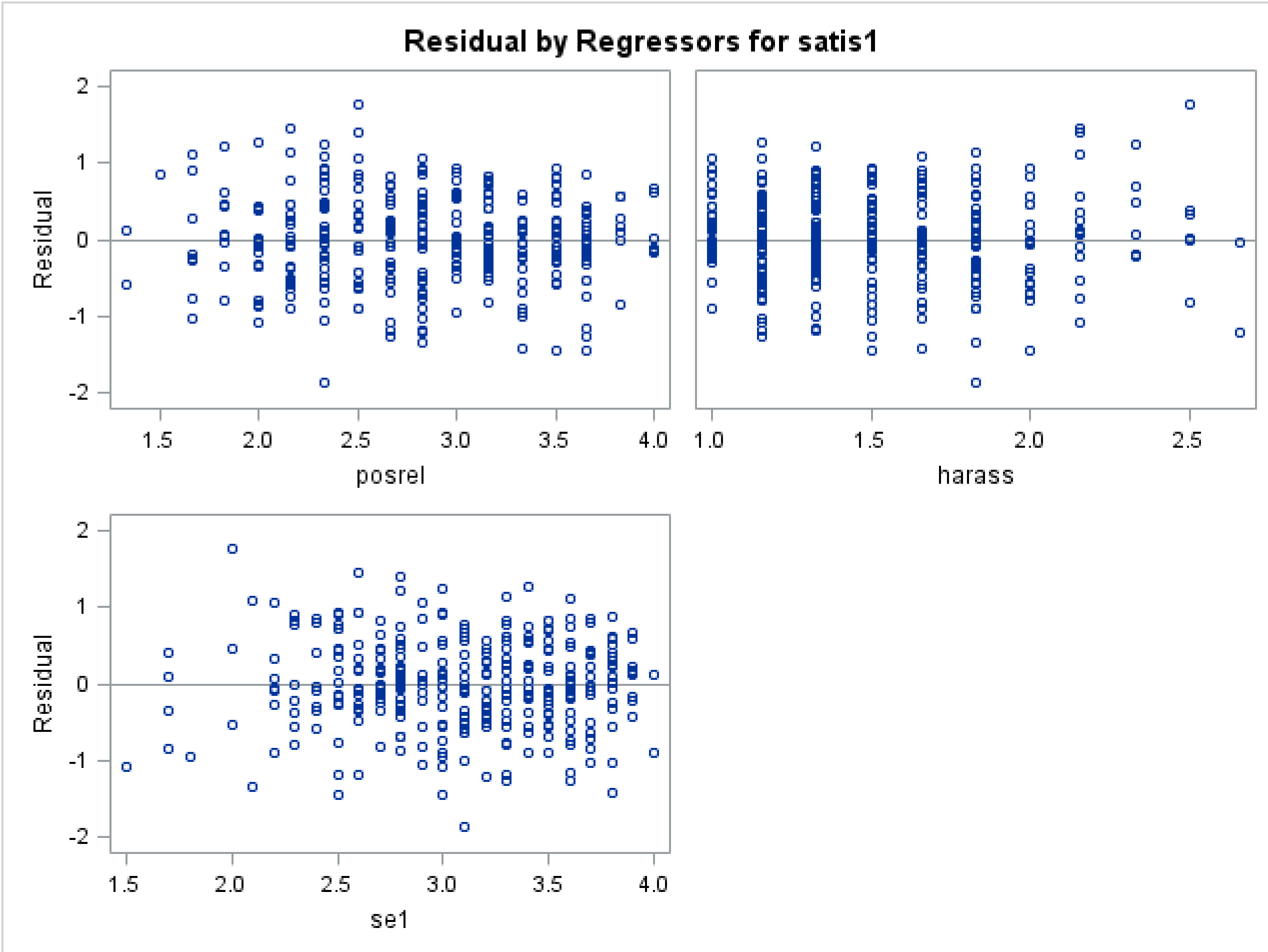
Root MSE	0.58503	**R-Square**	0.2969
Dependent Mean	3.75939	**Adj R-Sq**	0.2904
Coeff Var	15.56179		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate	95% Confidence Limits	
Intercept	**1**	1.47888	0.30890	4.79	<.0001	0	0.87119	2.08658
posrel	**1**	0.47230	0.05544	8.52	<.0001	0.40752	0.36323	0.58137
harass	**1**	-0.14514	0.08997	-1.61	0.1077	-0.07784	-0.32212	0.03185
se1	**1**	0.37292	0.06399	5.83	<.0001	0.27588	0.24703	0.49881

The REG Procedure
Model: MODEL1
Dependent Variable: satis1





Residual by Regressors for satis1
Residual
posrel
harass
se1
2
1
0
-1
-2
1.5
2.0
2.5
3.0
3.5
4.0
1.0

OUTPUT B: SPSS

```
regression/statistics defaults ci/dep=satis1/method=enter posrel harass se1.
```

Regression

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	se1, posrel, harass[b]	.	Enter

a. Dependent Variable: satis1

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.545[a]	.297	.290	.58503

a. Predictors: (Constant), se1, posrel, harass

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	47.120	3	15.707	45.891	.000[b]
	Residual	111.576	326	.342		
	Total	158.696	329			

a. Dependent Variable: satis1

b. Predictors: (Constant), se1, posrel, harass

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.479	.309		4.788	.000
	posrel	.472	.055	.408	8.519	.000
	harass	-.145	.090	-.078	-1.613	.108
	se1	.373	.064	.276	5.828	.000

Coefficients[a]

Model		95.0% Confidence Interval for B	
		Lower Bound	Upper Bound
1	(Constant)	.871	2.087
	posrel	.363	.581
	harass	-.322	.032
	se1	.247	.499

a. Dependent Variable: satis1

OUTPUT C: SAS

************************** PROCESS Procedure for SAS Version 3.0 ***************

Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

Documentation available in Hayes (2018) www.guilford.com/p/hayes3

****

Model and Variables

Model:	4
Y:	FAIL2
X:	HARASS
M:	SE2

Sample size:

330

****

OUTCOME VARIABLE:

SE2

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.2764	0.0764	0.2905	27.1349	1.0000	328.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	3.5966	0.1235	29.1227	0.0000	3.3536	3.8395
HARASS	-0.4156	0.0798	-5.2091	0.0000	-0.5725	-0.2586

****

OUTCOME VARIABLE:

FAIL2

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.3397	0.1154	0.2215	21.3247	2.0000	327.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	2.3845	0.2042	11.6757	0.0000	1.9827	2.7863
HARASS	0.0616	0.0725	0.8499	0.3960	-0.0810	0.2042
SE2	-0.2887	0.0482	-5.9879	0.0000	-0.3836	-0.1939

******************* TOTAL EFFECT MODEL ****************

OUTCOME VARIABLE:

FAIL2

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.1356	0.0184	0.2451	6.1419	1.0000	328.0000	0.0137

Model

	coeff	se	t	p	LLCI	ULCI
constant	1.3460	0.1134	11.8660	0.0000	1.1229	1.5692
HARASS	0.1816	0.0733	2.4783	0.0137	0.0374	0.3257

***************** TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y *******

Total effect of X on Y

Effect	se	t	p	LLCI	ULCI
0.1816	0.0733	2.4783	0.0137	0.0374	0.3257

Direct effect of X on Y

Effect	se	t	p	LLCI	ULCI
0.0616	0.0725	0.8499	0.3960	-0.0810	0.2042

Indirect effect(s) of X on Y:

	Effect	BootSE	BootLLCI	BootULCI
SE2	0.1200	0.0323	0.0629	0.1897

Normal theory test for indirect effect(s):

	Effect	se	Z	p
SE2	0.1200	0.0308	3.8993	0.0001

****************** ANALYSIS NOTES AND ERRORS ***************

Level of confidence for all confidence intervals in output:	
	95.0000
Number of bootstrap samples for percentile bootstrap confidence intervals:	
	10000

OUTPUT C: SPSS

```
Run MATRIX procedure:

**************** PROCESS Procedure for SPSS Version 3.00 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 4
    Y  : fail2
    X  : harass
    M  : se2

Sample
Size:  330

**************************************************************************
OUTCOME VARIABLE:
 se2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .2764      .0764      .2905    27.1349     1.0000   328.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.5966      .1235    29.1227      .0000     3.3536     3.8395
harass       -.4156      .0798    -5.2091      .0000     -.5725     -.2586

**************************************************************************
OUTCOME VARIABLE:
 fail2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .3397      .1154      .2215    21.3247     2.0000   327.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     2.3845      .2042    11.6757      .0000     1.9827     2.7863
harass        .0616      .0725      .8499      .3960     -.0810      .2042
se2          -.2887      .0482    -5.9879      .0000     -.3836     -.1939

************************** TOTAL EFFECT MODEL ****************************
OUTCOME VARIABLE:
 fail2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .1356      .0184      .2451     6.1419     1.0000   328.0000      .0137

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.3460      .1134    11.8660      .0000     1.1229     1.5692
```

```
harass        .1816      .0733     2.4783      .0137      .0374      .3257

************** TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Total effect of X on Y
     Effect         se          t          p       LLCI       ULCI
      .1816      .0733     2.4783      .0137      .0374      .3257

Direct effect of X on Y
     Effect         se          t          p       LLCI       ULCI
      .0616      .0725      .8499      .3960     -.0810      .2042

Indirect effect(s) of X on Y:
        Effect     BootSE   BootLLCI   BootULCI
se2      .1200      .0325      .0629      .1899

   Normal theory test for indirect effect(s):
        Effect         se          Z          p
se2      .1200      .0308     3.8993      .0001

*********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  10000

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

OUTPUT D: SAS

******************** PROCESS Procedure for SAS Version 3.0 ********************

Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

Documentation available in Hayes (2018) www.guilford.com/p/hayes3

****

Model and Variables

Model: 4

Y: FAIL2

X: HARASS

M: SE2

Covariates:

SE1 FAIL1

Sample size:

330

****

OUTCOME VARIABLE:

SE2

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.5450	0.2971	0.2224	45.9259	3.0000	326.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	2.0280	0.2412	8.4095	0.0000	1.5536	2.5024
HARASS	-0.2728	0.0717	-3.8025	0.0002	-0.4139	-0.1317

Model

	coeff	se	t	p	LLCI	ULCI
SE1	0.4879	0.0536	9.1081	0.0000	0.3825	0.5933
FAIL1	-0.1010	0.0606	-1.6661	0.0967	-0.2202	0.0182

****

OUTCOME VARIABLE:

FAIL2

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.4059	0.1648	0.2105	16.0276	4.0000	325.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	2.0934	0.2588	8.0900	0.0000	1.5843	2.6025
HARASS	0.0196	0.0713	0.2754	0.7832	-0.1207	0.1599
SE2	-0.2175	0.0539	-4.0375	0.0001	-0.3235	-0.1115
SE1	-0.0672	0.0584	-1.1517	0.2503	-0.1820	0.0476
FAIL1	0.2307	0.0592	3.8966	0.0001	0.1142	0.3471

************************ TOTAL EFFECT MODEL **************************

OUTCOME VARIABLE:

FAIL2

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.3505	0.1229	0.2203	15.2219	3.0000	326.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	1.6523	0.2400	6.8842	0.0000	1.1801	2.1245

Model

	coeff	se	t	p	LLCI	ULCI
HARASS	0.0790	0.0714	1.1060	0.2695	-0.0615	0.2194
SE1	-0.1733	0.0533	-3.2513	0.0013	-0.2782	-0.0685
FAIL1	0.2526	0.0603	4.1886	0.0000	0.1340	0.3713

********** TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y ***********

Total effect of X on Y

Effect	se	t	p	LLCI	ULCI
0.0790	0.0714	1.1060	0.2695	-0.0615	0.2194

Direct effect of X on Y

Effect	se	t	p	LLCI	ULCI
0.0196	0.0713	0.2754	0.7832	-0.1207	0.1599

Indirect effect(s) of X on Y:

	Effect	BootSE	BootLLCI	BootULCI
SE2	0.0593	0.0225	0.0207	0.1092

********************** ANALYSIS NOTES AND ERRORS ***********************

Level of confidence for all confidence intervals in output:
95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
10000

OUTPUT D: SPSS

```
Run MATRIX procedure:

**************** PROCESS Procedure for SPSS Version 3.00 ****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 4
    Y  : fail2
    X  : harass
    M  : se2

Covariates:
 se1      fail1

Sample
Size:  330

**************************************************************************
OUTCOME VARIABLE:
 se2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5450      .2971      .2224    45.9259     3.0000   326.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     2.0280      .2412     8.4095      .0000     1.5536     2.5024
harass       -.2728      .0717    -3.8025      .0002     -.4139     -.1317
se1           .4879      .0536     9.1081      .0000      .3825      .5933
fail1        -.1010      .0606    -1.6661      .0967     -.2202      .0182

**************************************************************************
OUTCOME VARIABLE:
 fail2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .4059      .1648      .2105    16.0276     4.0000   325.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     2.0934      .2588     8.0900      .0000     1.5843     2.6025
harass        .0196      .0713      .2754      .7832     -.1207      .1599
se2          -.2175      .0539    -4.0375      .0001     -.3235     -.1115
se1          -.0672      .0584    -1.1517      .2503     -.1820      .0476
fail1         .2307      .0592     3.8966      .0001      .1142      .3471

************************** TOTAL EFFECT MODEL ****************************
OUTCOME VARIABLE:
 fail2
```

Model Summary

R	R-sq	MSE	F	df1	df2	p
.3505	.1229	.2203	15.2219	3.0000	326.0000	.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	1.6523	.2400	6.8842	.0000	1.1801	2.1245
harass	.0790	.0714	1.1060	.2695	-.0615	.2194
se1	-.1733	.0533	-3.2513	.0013	-.2782	-.0685
fail1	.2526	.0603	4.1886	.0000	.1340	.3713

************** TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Total effect of X on Y

Effect	se	t	p	LLCI	ULCI
.0790	.0714	1.1060	.2695	-.0615	.2194

Direct effect of X on Y

Effect	se	t	p	LLCI	ULCI
.0196	.0713	.2754	.7832	-.1207	.1599

Indirect effect(s) of X on Y:

	Effect	BootSE	BootLLCI	BootULCI
se2	.0593	.0228	.0203	.1089

*********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
10000

------ END MATRIX -----

OUTPUT E: SAS

******************* PROCESS Procedure for SAS Version 3.0 ********************

Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

Documentation available in Hayes (2018) www.guilford.com/p/hayes3

****

Model and Variables

Model:	4
Y:	REACTION
X:	COND
M1:	IMPORT
M2:	PMI

Sample size:

123

****

OUTCOME VARIABLE:

IMPORT

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.1809	0.0327	2.9411	4.0942	1.0000	121.0000	0.0452

Model

	coeff	se	t	p	LLCI	ULCI
constant	3.9077	0.2127	18.3704	0.0000	3.4866	4.3288
COND	0.6268	0.3098	2.0234	0.0452	0.0135	1.2401

****

OUTCOME VARIABLE:

PMI

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.1808	0.0327	1.7026	4.0878	1.0000	121.0000	0.0454

Model

	coeff	se	t	p	LLCI	ULCI
constant	5.3769	0.1618	33.2222	0.0000	5.0565	5.6973
COND	0.4765	0.2357	2.0218	0.0454	0.0099	0.9431

****

OUTCOME VARIABLE:

REACTION

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.5702	0.3251	1.6628	19.1118	3.0000	119.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	-0.1498	0.5298	-0.2828	0.7778	-1.1989	0.8993
COND	0.1034	0.2391	0.4324	0.6662	-0.3701	0.5768
IMPORT	0.3244	0.0707	4.5857	0.0000	0.1843	0.4645
PMI	0.3965	0.0930	4.2645	0.0000	0.2124	0.5806

************************** TOTAL EFFECT MODEL ****************************

OUTCOME VARIABLE:

REACTION

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.1603	0.0257	2.3610	3.1897	1.0000	121.0000	0.0766

Model

	coeff	se	t	p	LLCI	ULCI
constant	3.2500	0.1906	17.0525	0.0000	2.8727	3.6273
COND	0.4957	0.2775	1.7860	0.0766	-0.0538	1.0452

************* TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Total effect of X on Y

Effect	se	t	p	LLCI	ULCI
0.4957	0.2775	1.7860	0.0766	-0.0538	1.0452

Direct effect of X on Y

Effect	se	t	p	LLCI	ULCI
0.1034	0.2391	0.4324	0.6662	-0.3701	0.5768

Indirect effect(s) of X on Y:

	Effect	BootSE	BootLLCI	BootULCI
TOTAL	0.3923	0.1647	0.0814	0.7304
IMPORT	0.2033	0.1158	0.0043	0.4613
PMI	0.1890	0.1037	0.0076	0.4143
(C1)	0.0144	0.1456	-0.2595	0.3104

Normal theory test for indirect effect(s):

	Effect	se	Z	p
IMPORT	0.2033	0.1120	1.8154	0.0695
PMI	0.1890	0.1057	1.7872	0.0739

Specific indirect effect contrast definition(s):

(C1)	IMPORT	minus	PMI

********************** ANALYSIS NOTES AND ERRORS ***********************

Level of confidence for all confidence intervals in output:	
95.0000	
Number of bootstrap samples for percentile bootstrap confidence intervals:	
10000	

Output E: SPSS

```
Run MATRIX procedure:

**************** PROCESS Procedure for SPSS Version 3.00 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 4
    Y  : react
    X  : cond
   M1  : import
   M2  : pmi

Sample
Size:  123

**************************************************************************
OUTCOME VARIABLE:
 import

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .1809      .0327     2.9411     4.0942     1.0000   121.0000      .0452

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.9077      .2127    18.3704      .0000     3.4866     4.3288
cond          .6268      .3098     2.0234      .0452      .0135     1.2401

**************************************************************************
OUTCOME VARIABLE:
 pmi

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .1808      .0327     1.7026     4.0878     1.0000   121.0000      .0454

Model
              coeff         se          t          p       LLCI       ULCI
constant     5.3769      .1618    33.2222      .0000     5.0565     5.6973
cond          .4765      .2357     2.0218      .0454      .0099      .9431

**************************************************************************
OUTCOME VARIABLE:
 react

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5702      .3251     1.6628    19.1118     3.0000   119.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     -.1498      .5298     -.2828      .7778    -1.1989      .8993
```

```
cond           .1034      .2391      .4324      .6662     -.3701      .5768
import         .3244      .0707     4.5857      .0000      .1843      .4645
pmi            .3965      .0930     4.2645      .0000      .2124      .5806

************************* TOTAL EFFECT MODEL ****************************
OUTCOME VARIABLE:
 react

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .1603      .0257     2.3610     3.1897     1.0000   121.0000      .0766

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.2500      .1906    17.0525      .0000     2.8727     3.6273
cond          .4957      .2775     1.7860      .0766     -.0538     1.0452

************** TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Total effect of X on Y
     Effect         se          t          p       LLCI       ULCI
      .4957      .2775     1.7860      .0766     -.0538     1.0452

Direct effect of X on Y
     Effect         se          t          p       LLCI       ULCI
      .1034      .2391      .4324      .6662     -.3701      .5768

Indirect effect(s) of X on Y:
           Effect     BootSE   BootLLCI   BootULCI
TOTAL        .3923      .1678      .0844      .7462
import       .2033      .1163      .0063      .4608
pmi          .1890      .1043      .0059      .4170
(C1)         .0144      .1437     -.2616      .3045

   Normal theory test for indirect effect(s):
           Effect         se          Z          p
import       .2033      .1120     1.8154      .0695
pmi          .1890      .1057     1.7872      .0739

Specific indirect effect contrast definition(s):
(C1)          import    minus   pmi

********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  10000

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

OUTPUT F: SAS

******************* PROCESS Procedure for SAS Version 3.0 ********************

Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

Documentation available in Hayes (2018) www.guilford.com/p/hayes3

****

Model and Variables

Model: 6

Y: MCIVIL

X: BINLADEN

M1: STEREO

M2: RTHREAT

Covariates:

SEX AGE IDEO

Sample size:

661

****

OUTCOME VARIABLE:

STEREO

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.3557	0.1265	0.6495	23.7609	4.0000	656.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	1.9045	0.1322	14.4084	0.0000	1.6449	2.1640

Model

	coeff	se	t	p	LLCI	ULCI
BINLADEN	0.1358	0.0639	2.1258	0.0339	0.0104	0.2613
SEX	0.0398	0.0635	0.6262	0.5314	-0.0849	0.1644
AGE	0.0504	0.0192	2.6220	0.0089	0.0127	0.0882
IDEO	0.1293	0.0143	9.0483	0.0000	0.1012	0.1574

****

OUTCOME VARIABLE:

RTHREAT

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.6764	0.4575	0.6076	110.4916	5.0000	655.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	-0.2548	0.1467	-1.7369	0.0829	-0.5428	0.0332
BINLADEN	0.0374	0.0620	0.6038	0.5462	-0.0843	0.1592
STEREO	0.7047	0.0378	18.6630	0.0000	0.6306	0.7789
SEX	0.1286	0.0614	2.0938	0.0367	0.0080	0.2492
AGE	0.0451	0.0187	2.4135	0.0161	0.0084	0.0818
IDEO	0.0898	0.0147	6.1257	0.0000	0.0610	0.1186

****

OUTCOME VARIABLE:

MCIVIL

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.6727	0.4526	0.5890	90.1100	6.0000	654.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	0.7165	0.1448	4.9499	0.0000	0.4323	1.0008
BINLADEN	-0.0311	0.0611	-0.5095	0.6106	-0.1510	0.0888
STEREO	0.1057	0.0460	2.2965	0.0220	0.0153	0.1960
RTHREAT	0.5491	0.0385	14.2732	0.0000	0.4736	0.6247
SEX	-0.1001	0.0607	-1.6504	0.0993	-0.2193	0.0190
AGE	-0.0103	0.0185	-0.5599	0.5758	-0.0466	0.0259
IDEO	0.0545	0.0148	3.6696	0.0003	0.0253	0.0836

************************* TOTAL EFFECT MODEL ***************************

OUTCOME VARIABLE:

MCIVIL

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.3675	0.1351	0.9278	25.6100	4.0000	656.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	1.5149	0.1580	9.5894	0.0000	1.2047	1.8251
BINLADEN	0.0564	0.0764	0.7380	0.4608	-0.0936	0.2063
SEX	-0.0099	0.0759	-0.1310	0.8958	-0.1589	0.1391
AGE	0.0393	0.0230	1.7085	0.0880	-0.0059	0.0844
IDEO	0.1675	0.0171	9.8053	0.0000	0.1339	0.2010

************* TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Total effect of X on Y

Effect	se	t	p	LLCI	ULCI
0.0564	0.0764	0.7380	0.4608	-0.0936	0.2063

Direct effect of X on Y

Effect	se	t	p	LLCI	ULCI
-0.0311	0.0611	-0.5095	0.6106	-0.1510	0.0888

Indirect effect(s) of X on Y:

	Effect	BootSE	BootLLCI	BootULCI
TOTAL	0.0875	0.0465	-0.0015	0.1804
Ind1	0.0144	0.0099	-0.0003	0.0369
Ind2	0.0206	0.0338	-0.0434	0.0885
Ind3	0.0526	0.0250	0.0057	0.1033

Indirect effect key:

Ind1	BINLADEN	->	STEREO	->	MCIVIL			
Ind2	BINLADEN	->	RTHREAT	->	MCIVIL			
Ind3	BINLADEN	->	STEREO	->	RTHREAT	->	MCIVIL	

********************** ANALYSIS NOTES AND ERRORS ***********************

Level of confidence for all confidence intervals in output:

95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:

10000

Output F: SPSS

```
Run MATRIX procedure:

**************** PROCESS Procedure for SPSS Version 3.00 ****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 6
    Y  : mcivil
    X  : binladen
   M1  : stereo
   M2  : rthreat

Covariates:
 sex      age      ideo

Sample
Size:  661

**************************************************************************
OUTCOME VARIABLE:
 stereo

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .3557      .1265      .6495    23.7609     4.0000   656.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.9045      .1322    14.4084      .0000     1.6449     2.1640
binladen      .1358      .0639     2.1258      .0339      .0104      .2613
sex           .0398      .0635      .6262      .5314     -.0849      .1644
age           .0504      .0192     2.6220      .0089      .0127      .0882
ideo          .1293      .0143     9.0483      .0000      .1012      .1574

**************************************************************************
OUTCOME VARIABLE:
 rthreat

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .6764      .4575      .6076   110.4916     5.0000   655.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     -.2548      .1467    -1.7369      .0829     -.5428      .0332
binladen      .0374      .0620      .6038      .5462     -.0843      .1592
stereo        .7047      .0378    18.6630      .0000      .6306      .7789
sex           .1286      .0614     2.0938      .0367      .0080      .2492
age           .0451      .0187     2.4135      .0161      .0084      .0818
ideo          .0898      .0147     6.1257      .0000      .0610      .1186

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

```
OUTCOME VARIABLE:
 mcivil

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .6727      .4526      .5890    90.1100     6.0000   654.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant      .7165      .1448     4.9499      .0000      .4323     1.0008
binladen     -.0311      .0611     -.5095      .6106     -.1510      .0888
stereo        .1057      .0460     2.2965      .0220      .0153      .1960
rthreat       .5491      .0385    14.2732      .0000      .4736      .6247
sex          -.1001      .0607    -1.6504      .0993     -.2193      .0190
age          -.0103      .0185     -.5599      .5758     -.0466      .0259
ideo          .0545      .0148     3.6696      .0003      .0253      .0836

************************** TOTAL EFFECT MODEL ****************************
OUTCOME VARIABLE:
 mcivil

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .3675      .1351      .9278    25.6100     4.0000   656.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.5149      .1580     9.5894      .0000     1.2047     1.8251
binladen      .0564      .0764      .7380      .4608     -.0936      .2063
sex          -.0099      .0759     -.1310      .8958     -.1589      .1391
age           .0393      .0230     1.7085      .0880     -.0059      .0844
ideo          .1675      .0171     9.8053      .0000      .1339      .2010

************** TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Total effect of X on Y
     Effect         se          t          p       LLCI       ULCI
      .0564      .0764      .7380      .4608     -.0936      .2063

Direct effect of X on Y
     Effect         se          t          p       LLCI       ULCI
     -.0311      .0611     -.5095      .6106     -.1510      .0888

Indirect effect(s) of X on Y:
          Effect     BootSE   BootLLCI   BootULCI
TOTAL      .0875      .0460      .0005      .1798
Ind1       .0144      .0098     -.0002      .0373
Ind2       .0206      .0338     -.0446      .0895
Ind3       .0526      .0248      .0050      .1021

Indirect effect key:
Ind1 binladen    ->    stereo      ->    mcivil
Ind2 binladen    ->    rthreat     ->    mcivil
Ind3 binladen    ->    stereo      ->    rthreat     ->    mcivil

*********************** ANALYSIS NOTES AND ERRORS ************************
```

```
Level of confidence for all confidence intervals in output:
  95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  10000

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

OUTPUT G: SAS

The REG Procedure
Model: MODEL1
Dependent Variable: crave2

Number of Observations Read	168
Number of Observations Used	168

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	42.32144	8.46429	11.63	<.0001
Error	162	117.88427	0.72768		
Corrected Total	167	160.20571			

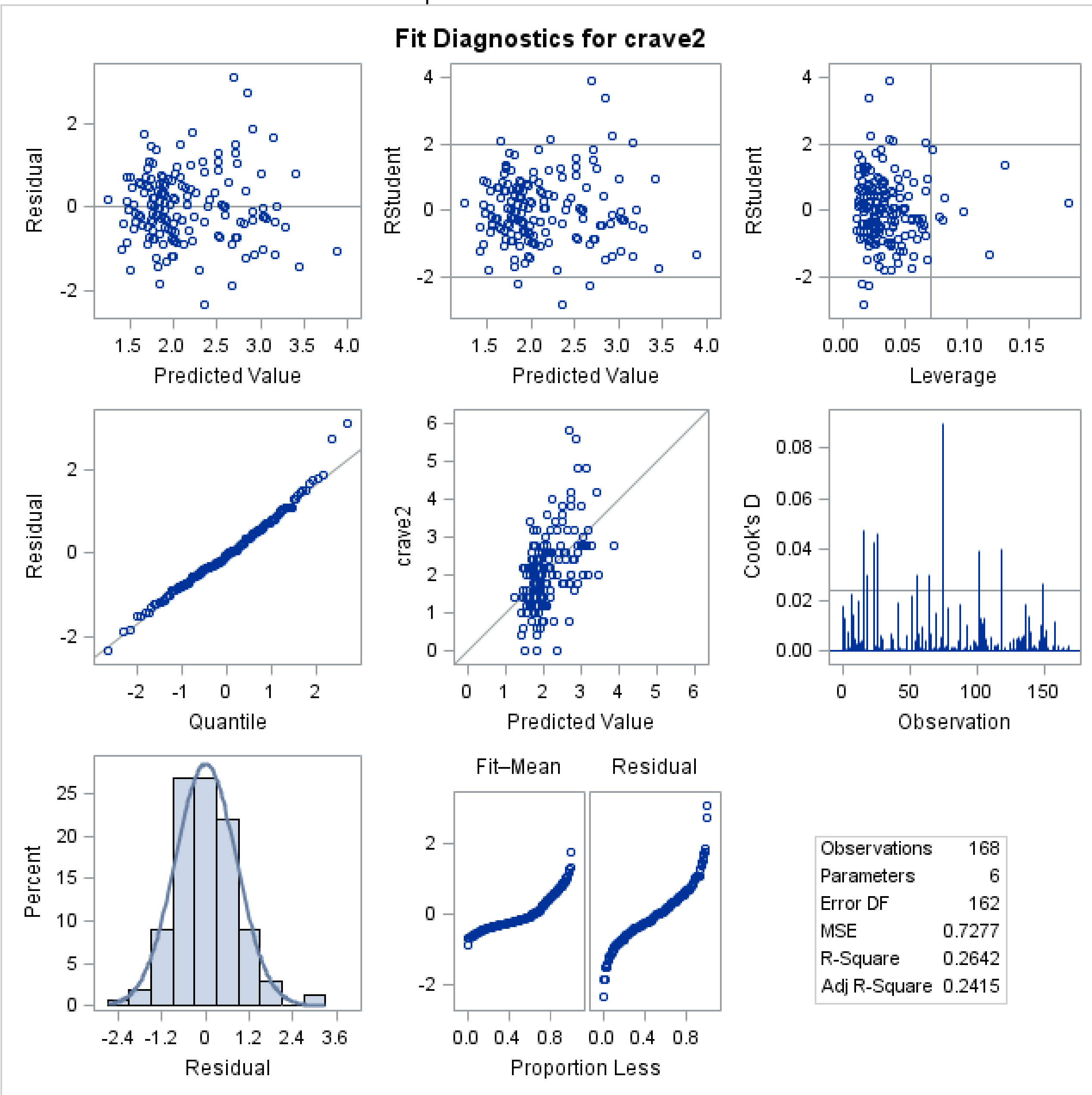
Root MSE	0.85304	**R-Square**	0.2642
Dependent Mean	2.11429	**Adj R-Sq**	0.2415
Coeff Var	40.34658		

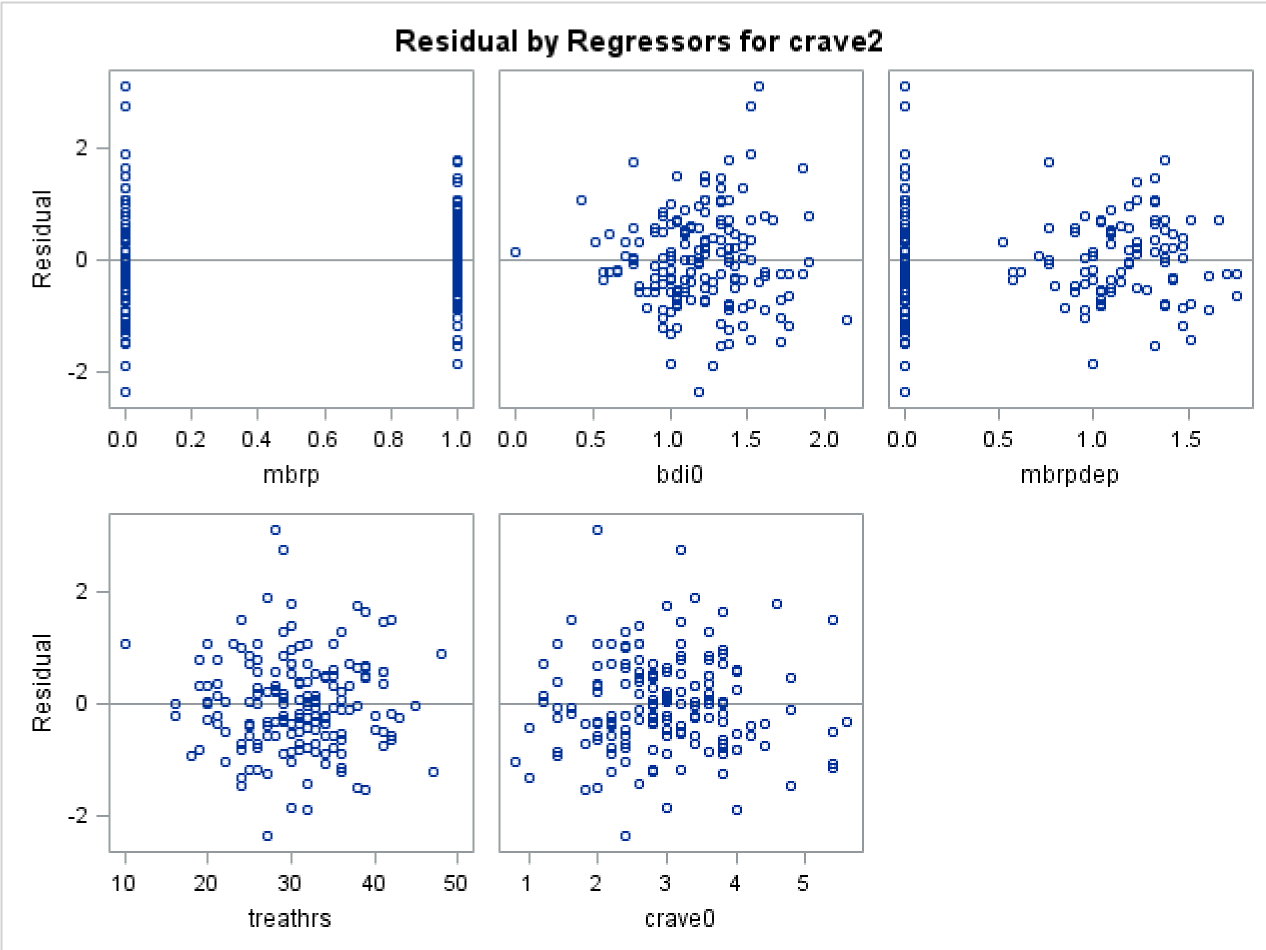
Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	**1**	1.03847	0.47010	2.21	0.0286
mbrp	**1**	0.58724	0.52413	1.12	0.2642
bdi0	**1**	1.12208	0.27620	4.06	<.0001
mbrpdep	**1**	-0.94845	0.42346	-2.24	0.0265
treathrs	**1**	-0.01767	0.01028	-1.72	0.0875
crave0	**1**	0.19204	0.07347	2.61	0.0098

The SAS System

The REG Procedure
Model: MODEL1
Dependent Variable: crave2




Residual by Regressors for crave2
Residual
2
0
-2
0.0
0.2
0.4
0.6
0.8
1.0
mbrp
0.0
0.5
1.0
1.5
2.0
bdi0
0.0
0.5
1.0
1.5
mbrpdep
Residual
2
0
-2
10
20
30
40
50
treathrs
1
2
3
4
5
crave0

OUTPUT G: SPSS

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	CRAVE0: Baseline craving, mbrpdep, TREATHRS: Hours of therapy, BDI0: Beck Depression Inventory baseline, MBRP: Therapy as usual (0) or MBRP therapy (1)[b]	.	Enter

a. Dependent Variable: CRAVE2: Craving at two month follow-up

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.514[a]	.264	.241	.8530

a. Predictors: (Constant), CRAVE0: Baseline craving, mbrpdep, TREATHRS: Hours of therapy, BDI0: Beck Depression Inventory baseline, MBRP: Therapy as usual (0) or MBRP therapy (1)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	42.321	5	8.464	11.632	.000[b]
	Residual	117.884	162	.728		
	Total	160.206	167			

a. Dependent Variable: CRAVE2: Craving at two month follow-up

b. Predictors: (Constant), CRAVE0: Baseline craving, mbrpdep, TREATHRS: Hours of therapy, BDI0: Beck Depression Inventory baseline, MBRP: Therapy as usual (0) or MBRP therapy (1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.038	.470		2.209	.029

	MBRP: Therapy as usual (0) or MBRP therapy (1)	.587	.524	.299	1.120	.264
	BDI0: Beck Depression Inventory baseline	1.122	.276	.366	4.063	.000
	mbrpdep	-.948	.423	-.598	-2.240	.026
	TREATHRS: Hours of therapy	-.018	.010	-.120	-1.719	.088
	CRAVE0: Baseline craving	.192	.073	.183	2.614	.010

a. Dependent Variable: CRAVE2: Craving at two month follow-up

OUTPUT H: SAS

******************* PROCESS Procedure for SAS Version 3.0 ********************

Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

Documentation available in Hayes (2018) www.guilford.com/p/hayes3

****

Model and Variables

Model:	1
Y:	CRAVE2
X:	MBRP
W:	BDI0

Covariates:

TREATHRS CRAVE0

Sample size:

168

****

OUTCOME VARIABLE:

CRAVE2

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.5140	0.2642	0.7277	11.6319	5.0000	162.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	1.0385	0.4701	2.2090	0.0286	0.1102	1.9668

Model

	coeff	se	t	p	LLCI	ULCI
MBRP	0.5872	0.5241	1.1204	0.2642	-0.4478	1.6222
BDI0	1.1221	0.2762	4.0625	0.0001	0.5767	1.6675
Int_1	-0.9485	0.4235	-2.2398	0.0265	-1.7847	-0.1122
TREATHRS	-0.0177	0.0103	-1.7190	0.0875	-0.0380	0.0026
CRAVE0	0.1920	0.0735	2.6138	0.0098	0.0470	0.3371

Product terms key:

Int_1 : MBRP x BDI0

Test(s) of highest order unconditional interactions:

	R2-chng	F	df1	df2	p
X*W	0.0228	5.0166	1.0000	162.0000	0.0265

Focal predict: MBRP (X)

Mod var: BDI0 (W)

Conditional effects of the focal predictor at values of the moderator(s):

BDI0	Effect	se	t	p	LLCI	ULCI
0.8772	-0.2447	0.1922	-1.2733	0.2047	-0.6243	0.1348
1.1962	-0.5473	0.1375	-3.9818	0.0001	-0.8188	-0.2759
1.5153	-0.8500	0.1933	-4.3973	0.0000	-1.2317	-0.4683

Moderator value(s) defining Johnson-Neyman significance region(s):

Value	% below	% above
0.9681	21.4286	78.5714

Conditional effect of focal predictor at values of the moderator:

BDI0	Effect	se	t	p	LLCI	ULCI
0.0000	0.5872	0.5241	1.1204	0.2642	-0.4478	1.6222
0.1070	0.4858	0.4806	1.0108	0.3136	-0.4632	1.4347
0.2140	0.3843	0.4373	0.8787	0.3809	-0.4793	1.2479
0.3210	0.2828	0.3946	0.7167	0.4746	-0.4964	1.0620
0.4280	0.1813	0.3525	0.5144	0.6077	-0.5147	0.8773
0.5350	0.0798	0.3112	0.2565	0.7979	-0.5348	0.6944
0.6420	-0.0217	0.2713	-0.0798	0.9365	-0.5574	0.5141
0.7490	-0.1231	0.2334	-0.5276	0.5985	-0.5840	0.3377
0.8560	-0.2246	0.1986	-1.1312	0.2596	-0.6167	0.1675
0.9630	-0.3261	0.1688	-1.9318	0.0551	-0.6595	0.0072
0.9681	-0.3309	0.1676	-1.9747	0.0500	-0.6618	0.0000
1.0700	-0.4276	0.1472	-2.9047	0.0042	-0.7183	-0.1369
1.1770	-0.5291	0.1377	-3.8435	0.0002	-0.8009	-0.2573
1.2840	-0.6306	0.1426	-4.4220	0.0000	-0.9122	-0.3490
1.3910	-0.7321	0.1607	-4.5553	0.0000	-1.0494	-0.4147
1.4980	-0.8335	0.1882	-4.4288	0.0000	-1.2052	-0.4619
1.6050	-0.9350	0.2216	-4.2186	0.0000	-1.3727	-0.4973
1.7120	-1.0365	0.2587	-4.0063	0.0001	-1.5474	-0.5256
1.8190	-1.1380	0.2981	-3.8178	0.0002	-1.7266	-0.5494
1.9260	-1.2395	0.3389	-3.6571	0.0003	-1.9088	-0.5702
2.0330	-1.3410	0.3808	-3.5216	0.0006	-2.0929	-0.5890
2.1400	-1.4424	0.4234	-3.4072	0.0008	-2.2784	-0.6064

Data for visualizing the conditional effect of the focal predictor:

MBRP	BDI0	CRAVE2
0.0000	0.8772	2.0456
1.0000	0.8772	1.8009
0.0000	1.1962	2.4037

MBRP	BDI0	CRAVE2
1.0000	1.1962	1.8563
0.0000	1.5153	2.7617
1.0000	1.5153	1.9117

********************** ANALYSIS NOTES AND ERRORS ***********************

Level of confidence for all confidence intervals in output:

95.0000

W values in conditional tables are the mean and +/- SD from the mean

OUTPUT H: SPSS

```
Run MATRIX procedure:

**************** PROCESS Procedure for SPSS Version 3.00 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 1
    Y  : crave2
    X  : mbrp
    W  : bdi0

Covariates:
 treathrs crave0

Sample
Size:  168

**************************************************************************
OUTCOME VARIABLE:
 crave2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5140      .2642      .7277    11.6319     5.0000   162.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.0385      .4701     2.2090      .0286      .1102     1.9668
mbrp          .5872      .5241     1.1204      .2642     -.4478     1.6222
bdi0         1.1221      .2762     4.0625      .0001      .5767     1.6675
Int_1        -.9485      .4235    -2.2398      .0265    -1.7847     -.1122
treathrs     -.0177      .0103    -1.7190      .0875     -.0380      .0026
crave0        .1920      .0735     2.6138      .0098      .0470      .3371

Product terms key:
 Int_1    :        mbrp     x        bdi0

Test(s) of highest order unconditional interaction(s):
       R2-chng          F        df1        df2          p
X*W      .0228     5.0166     1.0000   162.0000      .0265
----------
    Focal predict: mbrp     (X)
          Mod var: bdi0     (W)

Conditional effects of the focal predictor at values of the moderator(s):

       bdi0     Effect         se          t          p       LLCI       ULCI
      .8772     -.2447      .1922    -1.2733      .2047     -.6243      .1348
     1.1963     -.5473      .1375    -3.9818      .0001     -.8188     -.2759
     1.5153     -.8500      .1933    -4.3973      .0000    -1.2317     -.4683

Moderator value(s) defining Johnson-Neyman significance region(s):
```

Value	% below	% above
.9681	21.4286	78.5714

Conditional effect of focal predictor at values of the moderator:

bdi0	Effect	se	t	p	LLCI	ULCI
.0000	.5872	.5241	1.1204	.2642	-.4478	1.6222
.1070	.4858	.4806	1.0108	.3136	-.4632	1.4347
.2140	.3843	.4373	.8787	.3809	-.4793	1.2479
.3210	.2828	.3946	.7167	.4746	-.4964	1.0620
.4280	.1813	.3525	.5144	.6077	-.5147	.8773
.5350	.0798	.3112	.2565	.7979	-.5348	.6944
.6420	-.0217	.2713	-.0798	.9365	-.5574	.5141
.7490	-.1231	.2334	-.5276	.5985	-.5840	.3377
.8560	-.2246	.1986	-1.1312	.2596	-.6167	.1675
.9630	-.3261	.1688	-1.9318	.0551	-.6595	.0072
.9681	-.3309	.1676	-1.9747	.0500	-.6618	.0000
1.0700	-.4276	.1472	-2.9047	.0042	-.7183	-.1369
1.1770	-.5291	.1377	-3.8435	.0002	-.8009	-.2573
1.2840	-.6306	.1426	-4.4220	.0000	-.9122	-.3490
1.3910	-.7321	.1607	-4.5553	.0000	-1.0494	-.4147
1.4980	-.8335	.1882	-4.4288	.0000	-1.2052	-.4619
1.6050	-.9350	.2216	-4.2186	.0000	-1.3727	-.4973
1.7120	-1.0365	.2587	-4.0063	.0001	-1.5474	-.5256
1.8190	-1.1380	.2981	-3.8178	.0002	-1.7266	-.5494
1.9260	-1.2395	.3389	-3.6571	.0003	-1.9088	-.5702
2.0330	-1.3410	.3808	-3.5216	.0006	-2.0929	-.5890
2.1400	-1.4424	.4234	-3.4072	.0008	-2.2784	-.6064

Data for visualizing the conditional effect of the focal predictor:
Paste text below into a SPSS syntax window and execute to produce plot.

```
DATA LIST FREE/
   mbrp       bdi0       crave2     .
BEGIN DATA.
      .0000      .8772     2.0456
     1.0000      .8772     1.8009
      .0000     1.1963     2.4037
     1.0000     1.1963     1.8563
      .0000     1.5153     2.7617
     1.0000     1.5153     1.9117
END DATA.
GRAPH/SCATTERPLOT=
 bdi0     WITH     crave2   BY       mbrp     .
```

********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95.0000

W values in conditional tables are the mean and +/- SD from the mean.

------ END MATRIX -----

OUTPUT I: SAS

******************* PROCESS Procedure for SAS Version 3.0 ********************

Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

Documentation available in Hayes (2018) www.guilford.com/p/hayes3

****

Model and Variables

Model: 1

Y: CRAVE2

X: BDI0

W: MBRP

Covariates:

TREATHRS CRAVE0

Sample size:

168

****

OUTCOME VARIABLE:

CRAVE2

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.5140	0.2642	0.7277	11.6319	5.0000	162.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	1.0385	0.4701	2.2090	0.0286	0.1102	1.9668
BDI0	1.1221	0.2762	4.0625	0.0001	0.5767	1.6675

Model

	coeff	se	t	p	LLCI	ULCI
MBRP	0.5872	0.5241	1.1204	0.2642	-0.4478	1.6222
Int_1	-0.9485	0.4235	-2.2398	0.0265	-1.7847	-0.1122
TREATHRS	-0.0177	0.0103	-1.7190	0.0875	-0.0380	0.0026
CRAVE0	0.1920	0.0735	2.6138	0.0098	0.0470	0.3371

Product terms key:

Int_1 : BDI0 x MBRP

Test(s) of highest order unconditional interactions:

	R2-chng	F	df1	df2	p
X*W	0.0228	5.0166	1.0000	162.0000	0.0265

Focal predict: BDI0 (X)

Mod var: MBRP (W)

Conditional effects of the focal predictor at values of the moderator(s):

MBRP	Effect	se	t	p	LLCI	ULCI
0.0000	1.1221	0.2762	4.0625	0.0001	0.5767	1.6675
1.0000	0.1736	0.3281	0.5291	0.5974	-0.4744	0.8216

********************** ANALYSIS NOTES AND ERRORS ***********************

Level of confidence for all confidence intervals in output:

95.0000

OUTPUT I: SPSS

```
Run MATRIX procedure:

**************** PROCESS Procedure for SPSS Version 3.00 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 1
    Y  : crave2
    X  : bdi0
    W  : mbrp

Covariates:
 treathrs crave0

Sample
Size:  168

**************************************************************************
OUTCOME VARIABLE:
 crave2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5140      .2642      .7277    11.6319     5.0000   162.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.0385      .4701     2.2090      .0286      .1102     1.9668
bdi0         1.1221      .2762     4.0625      .0001      .5767     1.6675
mbrp          .5872      .5241     1.1204      .2642     -.4478     1.6222
Int_1        -.9485      .4235    -2.2398      .0265    -1.7847     -.1122
treathrs     -.0177      .0103    -1.7190      .0875     -.0380      .0026
crave0        .1920      .0735     2.6138      .0098      .0470      .3371

Product terms key:
 Int_1    :        bdi0     x        mbrp

Test(s) of highest order unconditional interaction(s):
       R2-chng          F        df1        df2          p
X*W      .0228     5.0166     1.0000   162.0000      .0265
----------
    Focal predict: bdi0     (X)
          Mod var: mbrp     (W)

Conditional effects of the focal predictor at values of the moderator(s):

       mbrp     Effect         se          t          p       LLCI       ULCI
      .0000     1.1221      .2762     4.0625      .0001      .5767     1.6675
     1.0000      .1736      .3281      .5291      .5974     -.4744      .8216

********************** ANALYSIS NOTES AND ERRORS ************************
```

```
Level of confidence for all confidence intervals in output:
  95.0000

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

OUTPUT J: SAS

******************** PROCESS Procedure for SAS Version 3.0 ********************

Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

Documentation available in Hayes (2018) www.guilford.com/p/hayes3

****

Model and Variables

Model: 14

Y: PERFORM

X: DYSFUNC

M: NEGTONE

W: NEGEXP

Covariates:

D1 D2 D3

Sample size:

60

****

OUTCOME VARIABLE:

NEGTONE

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.5026	0.2526	0.2213	4.6462	4.0000	55.0000	0.0027

Model

	coeff	se	t	p	LLCI	ULCI
constant	-0.2057	0.1305	-1.5760	0.1208	-0.4672	0.0559

Model

	coeff	se	t	p	LLCI	ULCI
DYSFUNC	0.6095	0.1668	3.6546	0.0006	0.2753	0.9437
D1	0.3487	0.1715	2.0332	0.0469	0.0050	0.6923
D2	0.2951	0.2122	1.3906	0.1700	-0.1302	0.7204
D3	0.2507	0.1663	1.5078	0.1373	-0.0825	0.5840

****

OUTCOME VARIABLE:

PERFORM

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.5937	0.3524	0.2006	4.0428	7.0000	52.0000	0.0013

Model

	coeff	se	t	p	LLCI	ULCI
constant	-0.1754	0.1305	-1.3444	0.1847	-0.4373	0.0864
DYSFUNC	0.3729	0.1808	2.0622	0.0442	0.0100	0.7357
NEGTONE	-0.4886	0.1377	-3.5485	0.0008	-0.7649	-0.2123
NEGEXP	-0.0221	0.1176	-0.1875	0.8520	-0.2581	0.2140
Int_1	-0.4498	0.2451	-1.8353	0.0722	-0.9417	0.0420
D1	0.1815	0.1720	1.0556	0.2960	-0.1635	0.5266
D2	0.0841	0.2099	0.4004	0.6905	-0.3372	0.5053
D3	0.2816	0.1648	1.7087	0.0935	-0.0491	0.6123

Product terms key:

Int_1 : NEGTONE x NEGEXP

Test(s) of highest order unconditional interactions:

	R2-chng	F	df1	df2	p
M*W	0.0419	3.3684	1.0000	52.0000	0.0722

Focal predict: NEGTONE (M)

Mod var: NEGEXP (W)

Conditional effects of the focal predictor at values of the moderator(s):

NEGEXP	**Effect**	**se**	**t**	**p**	**LLCI**	**ULCI**
-0.5308	-0.2498	0.2196	-1.1379	0.2604	-0.6904	0.1907
-0.0600	-0.4616	0.1434	-3.2188	0.0022	-0.7494	-0.1738
0.6000	-0.7585	0.1633	-4.6451	0.0000	-1.0862	-0.4308

Data for visualizing the conditional effect of the focal predictor:

NEGTONE	**NEGEXP**	**PERFORM**
-0.4500	-0.5308	0.1258
-0.0350	-0.5308	0.0222
0.5224	-0.5308	-0.1171
-0.4500	-0.0600	0.2108
-0.0350	-0.0600	0.0192
0.5224	-0.0600	-0.2381
-0.4500	0.6000	0.3298
-0.0350	0.6000	0.0150
0.5224	0.6000	-0.4078

**************** DIRECT AND INDIRECT EFFECTS OF X ON Y *****************

Direct effect of X on Y

Effect	**se**	**t**	**p**	**LLCI**	**ULCI**
0.3729	0.1808	2.0622	0.0442	0.0100	0.7357

Conditional indirect effects of X on Y:

INDIRECT EFFECT:

DYSFUNC -> NEGTONE -> PERFORM

NEGEXP	**Effect**	**BootSE**	**BootLLCI**	**BootULCI**
-0.5308	-0.1523	0.1517	-0.4393	0.1830
-0.0600	-0.2813	0.1250	-0.5451	-0.0571
0.6000	-0.4623	0.1690	-0.8083	-0.1472

Index of moderated mediation:

	Index	**BootSE**	**BootLLCI**	**BootULCI**
NEGEXP	-0.2742	0.1768	-0.7025	-0.0234

*********************** ANALYSIS NOTES AND ERRORS ***********************

Level of confidence for all confidence intervals in output:

95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:

10000

W values in conditional tables are the 16th, 50th, and 84th percentiles

NOTE: Some bootstrap samples had to be replaced. The number of such replacements was:

OUTPUT J: SPSS

Run MATRIX procedure:

**************** PROCESS Procedure for SPSS Version 3.00 ****************

Written by Andrew F. Hayes, Ph.D. www.afhayes.com
Documentation available in Hayes (2018). www.guilford.com/p/hayes3

****
Model : 14
Y : perform
X : dysfunc
M : negtone
W : negexp

Covariates:
d1 d2 d3

Sample
Size: 60

****
OUTCOME VARIABLE:
negtone

Model Summary

R	R-sq	MSE	F	df1	df2	p
.5026	.2526	.2213	4.6462	4.0000	55.0000	.0027

Model

	coeff	se	t	p	LLCI	ULCI
constant	-.2057	.1305	-1.5760	.1208	-.4672	.0559
dysfunc	.6095	.1668	3.6546	.0006	.2753	.9437
d1	.3487	.1715	2.0332	.0469	.0050	.6923
d2	.2951	.2122	1.3906	.1700	-.1302	.7204
d3	.2507	.1663	1.5078	.1373	-.0825	.5840

****
OUTCOME VARIABLE:
perform

Model Summary

R	R-sq	MSE	F	df1	df2	p
.5937	.3524	.2006	4.0428	7.0000	52.0000	.0013

Model

	coeff	se	t	p	LLCI	ULCI
constant	-.1754	.1305	-1.3444	.1847	-.4373	.0864
dysfunc	.3729	.1808	2.0622	.0442	.0100	.7357
negtone	-.4886	.1377	-3.5485	.0008	-.7649	-.2123
negexp	-.0221	.1176	-.1875	.8520	-.2581	.2140
Int_1	-.4498	.2451	-1.8353	.0722	-.9417	.0420
d1	.1815	.1720	1.0556	.2960	-.1635	.5266
d2	.0841	.2099	.4004	.6905	-.3372	.5053
d3	.2816	.1648	1.7087	.0935	-.0491	.6123

```
Product terms key:
 Int_1    :        negtone  x        negexp

Test(s) of highest order unconditional interaction(s):
       R2-chng          F        df1        df2          p
M*W      .0419     3.3684     1.0000    52.0000      .0722
----------
    Focal predict: negtone  (M)
          Mod var: negexp   (W)

Conditional effects of the focal predictor at values of the moderator(s):

     negexp     Effect         se          t          p       LLCI       ULCI
     -.5308     -.2498      .2196    -1.1379      .2604     -.6904      .1907
     -.0600     -.4616      .1434    -3.2188      .0022     -.7494     -.1738
      .6000     -.7585      .1633    -4.6451      .0000    -1.0862     -.4308

****************** DIRECT AND INDIRECT EFFECTS OF X ON Y *****************

Direct effect of X on Y
     Effect         se          t          p       LLCI       ULCI
      .3729      .1808     2.0622      .0442      .0100      .7357

Conditional indirect effects of X on Y:

INDIRECT EFFECT:
 dysfunc     ->    negtone     ->    perform

     negexp     Effect     BootSE   BootLLCI   BootULCI
     -.5308     -.1523      .1497     -.4365      .1726
     -.0600     -.2813      .1249     -.5472     -.0569
      .6000     -.4623      .1683     -.8113     -.1543

      Index of moderated mediation:
            Index     BootSE   BootLLCI   BootULCI
negexp     -.2742      .1727     -.6833     -.0243
---

********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  10000

W values in conditional tables are the 16th, 50th, and 84th percentiles.

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

OUTPUT K: SAS

*********************** PROCESS Procedure for SAS Version 3.0 ******************

Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

Documentation available in Hayes (2018) www.guilford.com/p/hayes3

****

Model and Variables

Model:	7
Y:	USE4
X:	MBRP
M:	CRAVE2
W:	BDI0

Covariates:

CRAVE0 TREATHRS

Sample size:

168

****

OUTCOME VARIABLE:

CRAVE2

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.5140	0.2642	0.7277	11.6319	5.0000	162.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI

Model

	coeff	se	t	p	LLCI	ULCI
constant	1.0385	0.4701	2.2090	0.0286	0.1102	1.9668
MBRP	0.5872	0.5241	1.1204	0.2642	-0.4478	1.6222
BDI0	1.1221	0.2762	4.0625	0.0001	0.5767	1.6675
Int_1	-0.9485	0.4235	-2.2398	0.0265	-1.7847	-0.1122
CRAVE0	0.1920	0.0735	2.6138	0.0098	0.0470	0.3371
TREATHRS	-0.0177	0.0103	-1.7190	0.0875	-0.0380	0.0026

Product terms key:

Int_1 : MBRP x BDI0

Test(s) of highest order unconditional interactions:

	R2-chng	F	df1	df2	p
X*W	0.0228	5.0166	1.0000	162.0000	0.0265

Focal predict: MBRP (X)

Mod var: BDI0 (W)

Conditional effects of the focal predictor at values of the moderator(s):

BDI0	Effect	se	t	p	LLCI	ULCI
0.9020	-0.2683	0.1850	-1.4500	0.1490	-0.6336	0.0971
1.1900	-0.5414	0.1375	-3.9384	0.0001	-0.8129	-0.2699
1.5180	-0.8525	0.1941	-4.3923	0.0000	-1.2358	-0.4692

****

OUTCOME VARIABLE:

USE4

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.7304	0.5335	0.2105	46.6070	4.0000	163.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	1.1298	0.2150	5.2545	0.0000	0.7052	1.5544
MBRP	0.0926	0.0773	1.1979	0.2327	-0.0601	0.2453
CRAVE2	0.4810	0.0402	11.9547	0.0000	0.4015	0.5604
CRAVE0	-0.0884	0.0397	-2.2246	0.0275	-0.1668	-0.0099
TREATHRS	-0.0199	0.0056	-3.5720	0.0005	-0.0309	-0.0089

*************** DIRECT AND INDIRECT EFFECTS OF X ON Y ******************

Direct effect of X on Y

Effect	se	t	p	LLCI	ULCI
0.0926	0.0773	1.1979	0.2327	-0.0601	0.2453

Conditional indirect effects of X on Y:

INDIRECT EFFECT:

MBRP -> CRAVE2 -> USE4

BDI0	Effect	BootSE	BootLLCI	BootULCI
0.9020	-0.1290	0.0756	-0.2858	0.0179
1.1900	-0.2604	0.0856	-0.4445	-0.1122
1.5180	-0.4100	0.1365	-0.7073	-0.1770

Index of moderated mediation:

	Index	BootSE	BootLLCI	BootULCI
BDI0	-0.4562	0.2162	-0.9372	-0.0890

******************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:

95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:

10000

W values in conditional tables are the 16th, 50th, and 84th percentiles

OUTPUT K: SPSS

```
Run MATRIX procedure:

**************** PROCESS Procedure for SPSS Version 3.00 ****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 7
    Y  : use4
    X  : mbrp
    M  : crave2
    W  : bdi0

Covariates:
 crave0   treathrs

Sample
Size:  168

**************************************************************************
OUTCOME VARIABLE:
 crave2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5140      .2642      .7277    11.6319     5.0000   162.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.0385      .4701     2.2090      .0286      .1102     1.9668
mbrp          .5872      .5241     1.1204      .2642     -.4478     1.6222
bdi0         1.1221      .2762     4.0625      .0001      .5767     1.6675
Int_1        -.9485      .4235    -2.2398      .0265    -1.7847     -.1122
crave0        .1920      .0735     2.6138      .0098      .0470      .3371
treathrs     -.0177      .0103    -1.7190      .0875     -.0380      .0026

Product terms key:
 Int_1    :        mbrp     x        bdi0

Test(s) of highest order unconditional interaction(s):
       R2-chng          F        df1        df2          p
X*W      .0228     5.0166     1.0000   162.0000      .0265
----------
    Focal predict: mbrp     (X)
          Mod var: bdi0     (W)

Conditional effects of the focal predictor at values of the moderator(s):

       bdi0     Effect         se          t          p       LLCI       ULCI
      .9020     -.2683      .1850    -1.4500      .1490     -.6336      .0971
     1.1900     -.5414      .1375    -3.9384      .0001     -.8129     -.2699
     1.5180     -.8525      .1941    -4.3923      .0000    -1.2358     -.4692
```

```
**************************************************************************
OUTCOME VARIABLE:
 use4

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .7304      .5335      .2105    46.6070     4.0000   163.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.1298      .2150     5.2545      .0000      .7052     1.5544
mbrp          .0926      .0773     1.1979      .2327     -.0601      .2453
crave2        .4810      .0402    11.9547      .0000      .4015      .5604
crave0       -.0884      .0397    -2.2246      .0275     -.1668     -.0099
treathrs     -.0199      .0056    -3.5720      .0005     -.0309     -.0089

***************** DIRECT AND INDIRECT EFFECTS OF X ON Y ****************

Direct effect of X on Y
     Effect         se          t          p       LLCI       ULCI
      .0926      .0773     1.1979      .2327     -.0601      .2453

Conditional indirect effects of X on Y:

INDIRECT EFFECT:
 mbrp        ->    crave2      ->    use4

       bdi0     Effect     BootSE   BootLLCI   BootULCI
      .9020     -.1290      .0770     -.2869      .0183
     1.1900     -.2604      .0862     -.4458     -.1090
     1.5180     -.4100      .1367     -.7080     -.1767

      Index of moderated mediation:
          Index     BootSE   BootLLCI   BootULCI
bdi0     -.4562      .2172     -.9463     -.0934
---

********************** ANALYSIS NOTES AND ERRORS ***********************

Level of confidence for all confidence intervals in output:
  95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  10000

W values in conditional tables are the 16th, 50th, and 84th percentiles.

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

OUTPUT L: SAS

********************* PROCESS Procedure for SAS Version 3.0 *********************

Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

Documentation available in Hayes (2018) www.guilford.com/p/hayes3

****

Model and Variables

Model:	4
Y:	LIKING
X:	COND
M:	RESPAPPR

Sample size:

129

Coding of categorical X variable for analysis:

COND	X1	X2
0	0	0
1	1	0
2	0	1

****

OUTCOME VARIABLE:

RESPAPPR

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.5106	0.2607	1.3649	22.2190	2.0000	126.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	3.8841	0.1825	21.2881	0.0000	3.5231	4.2452
X1	1.2612	0.2550	4.9456	0.0000	0.7565	1.7659
X2	1.6103	0.2522	6.3842	0.0000	1.1111	2.1095

****

OUTCOME VARIABLE:

LIKING

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.5031	0.2531	0.8427	14.1225	3.0000	125.0000	0.0000

Model

	coeff	se	t	p	LLCI	ULCI
constant	3.7103	0.3074	12.0711	0.0000	3.1020	4.3187
X1	-0.0037	0.2190	-0.0169	0.9865	-0.4371	0.4297
X2	-0.2202	0.2280	-0.9658	0.3360	-0.6715	0.2310
RESPAPPR	0.4119	0.0700	5.8844	0.0000	0.2734	0.5504

***************************** TOTAL EFFECT MODEL *********************

OUTCOME VARIABLE:

LIKING

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.2151	0.0463	1.0676	3.0552	2.0000	126.0000	0.0506

Model

	coeff	se	t	p	LLCI	ULCI
constant	5.3102	0.1614	32.9083	0.0000	4.9909	5.6296
X1	0.5158	0.2255	2.2870	0.0239	0.0695	0.9621

Model

	coeff	se	t	p	LLCI	ULCI
X2	0.4431	0.2231	1.9863	0.0492	0.0016	0.8845

*************** TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y ***********

Relative total effects of X on Y:

	Effect	se	t	p	LLCI	ULCI
X1	0.5158	0.2255	2.2870	0.0239	0.0695	0.9621
X2	0.4431	0.2231	1.9863	0.0492	0.0016	0.8845

Omnibus test of total effect of X on Y:

R2-chng	F	df1	df2	p
0.0463	3.0552	2.0000	126.0000	0.0506

Relative direct effects of X on Y

	Effect	se	t	p	LLCI	ULCI
X1	-0.0037	0.2190	-0.0169	0.9865	-0.4371	0.4297
X2	-0.2202	0.2280	-0.9658	0.3360	-0.6715	0.2310

Omnibus test of direct effect of X on Y:

R2-chng	F	df1	df2	p
0.0087	0.7286	2.0000	125.0000	0.4846

Relative indirect effects of X on Y

COND -> RESPAPPR -> LIKING

Effect BootSE BootLLCI BootULCI

	Effect	BootSE	BootLLCI	BootULCI
X1	0.5195	0.1509	0.2471	0.8401
X2	0.6633	0.1657	0.3593	1.0059

************************* ANALYSIS NOTES AND ERRORS ***********************

Level of confidence for all confidence intervals in output:

95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:

10000

OUTPUT L: SPSS

```
Run MATRIX procedure:

**************** PROCESS Procedure for SPSS Version 3.00 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 7
    Y  : use4
    X  : mbrp
    M  : crave2
    W  : bdi0

Covariates:
 crave0   treathrs

Sample
Size:  168

**************************************************************************
OUTCOME VARIABLE:
 crave2

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5140      .2642      .7277    11.6319     5.0000   162.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.0385      .4701     2.2090      .0286      .1102     1.9668
mbrp          .5872      .5241     1.1204      .2642     -.4478     1.6222
bdi0         1.1221      .2762     4.0625      .0001      .5767     1.6675
Int_1        -.9485      .4235    -2.2398      .0265    -1.7847     -.1122
crave0        .1920      .0735     2.6138      .0098      .0470      .3371
treathrs     -.0177      .0103    -1.7190      .0875     -.0380      .0026

Product terms key:
 Int_1    :        mbrp     x        bdi0

Test(s) of highest order unconditional interaction(s):
       R2-chng          F        df1        df2          p
X*W      .0228     5.0166     1.0000   162.0000      .0265
----------
    Focal predict: mbrp     (X)
          Mod var: bdi0     (W)

Conditional effects of the focal predictor at values of the moderator(s):

       bdi0     Effect         se          t          p       LLCI       ULCI
      .9020     -.2683      .1850    -1.4500      .1490     -.6336      .0971
     1.1900     -.5414      .1375    -3.9384      .0001     -.8129     -.2699
     1.5180     -.8525      .1941    -4.3923      .0000    -1.2358     -.4692
```

```
**************************************************************************
OUTCOME VARIABLE:
 use4

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .7304      .5335      .2105    46.6070     4.0000   163.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1.1298      .2150     5.2545      .0000      .7052     1.5544
mbrp          .0926      .0773     1.1979      .2327     -.0601      .2453
crave2        .4810      .0402    11.9547      .0000      .4015      .5604
crave0       -.0884      .0397    -2.2246      .0275     -.1668     -.0099
treathrs     -.0199      .0056    -3.5720      .0005     -.0309     -.0089

****************** DIRECT AND INDIRECT EFFECTS OF X ON Y *****************

Direct effect of X on Y
     Effect         se          t          p       LLCI       ULCI
      .0926      .0773     1.1979      .2327     -.0601      .2453

Conditional indirect effects of X on Y:

INDIRECT EFFECT:
 mbrp        ->    crave2      ->    use4

       bdi0     Effect     BootSE   BootLLCI   BootULCI
      .9020     -.1290      .0770     -.2869      .0183
     1.1900     -.2604      .0862     -.4458     -.1090
     1.5180     -.4100      .1367     -.7080     -.1767

      Index of moderated mediation:
          Index     BootSE   BootLLCI   BootULCI
bdi0     -.4562      .2172     -.9463     -.0934
---

*********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  10000

W values in conditional tables are the 16th, 50th, and 84th percentiles.

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

OUTPUT M: SAS

```
*************** MEMORE Procedure for SAS Version 1.1 ****************

          Written by Amanda K. Montoya and Andrew F. Hayes

               Documentation available at afhayes.com

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

                           Variables:
                      Y = PAIN2    PAIN1
                      M = HORMONE2 HORMONE1

                       Computed Variables:
            Ydiff =   PAIN2    - PAIN1
            Mdiff =   HORMONE2 - HORMONE1
            Mavg =  ( HORMONE2 + HORMONE1 ) /2 centered

                          Sample Size:
                               20

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

                            Outcome:
                      Ydiff = PAIN2 - PAIN1

                              Model
         Effect      SE        t       df       p      LLCI    ULCI
 'X' -6.5000 2.0654 -3.1471 19.0000 0.0053 -10.8232 -2.1768

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

                            Outcome:
                   Mdiff = HORMONE2 - HORMONE1

                              Model
```

	Effect	SE	t	df	p	LLCI	ULCI
'X'	-2.2500	0.9173	-2.4528	19.0000	0.0240	-4.1701	-0.3299

****

Outcome:

Ydiff = PAIN2 - PAIN1

Model Summary

R	R-sq	MSE	F	df1	df2	p
0.5554	0.3085	65.9384	3.7918	2.0000	17.0000	0.0435

Model

	Effect	SE	t	df	p	LLCI	ULCI
'X'	-3.7654	2.0869	-1.8043	17.0000	0.0889	-8.1687	0.6379
Mdiff	1.2154	0.4572	2.6583	17.0000	0.0166	0.2507	2.1801
Mavg	-0.1302	0.3209	-0.4057	17.0000	0.6900	-0.8074	0.5470

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

Total effect of X on Y

Effect	SE	t	df	p	LLCI	ULCI
-6.5000	2.0654	-3.1471	19.0000	0.0053	-10.8232	-2.1768

Direct effect of X on Y

Effect	SE	t	df	p	LLCI	ULCI
-3.7654	2.0869	-1.8043	17.0000	0.0889	-8.1687	0.6379

Indirect Effect of X on Y through M

	Effect	Boot SE	BootLLCI	BootULCI
Ind1	-2.7346	1.3262	-5.6979	-0.5390

Indirect effect key

Ind1 X -> M1diff -> Ydiff

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

Check SAS log for errors. Do not interpret output if errors are found.

Bootstrap confidence interval method: Percentile

Number of samples for bootstrap confidence intervals:
10000

Level of confidence for all confidence intervals in output:
95

OUTPUT M: SPSS

```
Run MATRIX procedure:

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

                           Written by Amanda Montoya

                    Documentation available at afhayes.com

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

Variables:
Y = pain2    pain1
M = hormone2 hormone1

Computed Variables:
Ydiff =          pain2     -       pain1
Mdiff =          hormone2  -       hormone1
Mavg =  (        hormone2  +       hormone1 )        /2       Centered

Sample Size:
  20

*****************************************************************************
Outcome: Ydiff =  pain2     -       pain1

Model
       Effect         SE          t         df          p       LLCI       ULCI
'X'    -6.5000     2.0654    -3.1471    19.0000      .0053   -10.8233    -2.1767

*****************************************************************************
Outcome: Mdiff =  hormone2  -       hormone1

Model
       Effect         SE          t         df          p       LLCI       ULCI
'X'    -2.2500      .9173    -2.4528    19.0000      .0240    -4.1701     -.3299

*****************************************************************************
Outcome: Ydiff =  pain2     -       pain1

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5554      .3085    65.9384     3.7918     2.0000    17.0000      .0435

Model
          coeff         SE          t         df          p       LLCI       ULCI
'X'     -3.7654     2.0869    -1.8043    17.0000      .0889    -8.1689      .6381
Mdiff    1.2154      .4572     2.6583    17.0000      .0166      .2506     2.1801
Mavg     -.1302      .3209     -.4057    17.0000      .6900     -.8074      .5470

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

Total effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
    -6.5000     2.0654    -3.1471    19.0000      .0053   -10.8233    -2.1767

Direct effect of X on Y
     Effect         SE          t         df          p       LLCI       ULCI
    -3.7654     2.0869    -1.8043    17.0000      .0889    -8.1689      .6381

Indirect Effect of X on Y through M
          Effect     BootSE   BootLLCI   BootULCI
```

```
Ind1     -2.7346     1.3069    -5.6645     -.5186

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

Level of confidence for all confidence intervals in output:
      95.00

The following variables were mean centered prior to analysis:
 (        hormone2  +       hormone1 )        /2

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