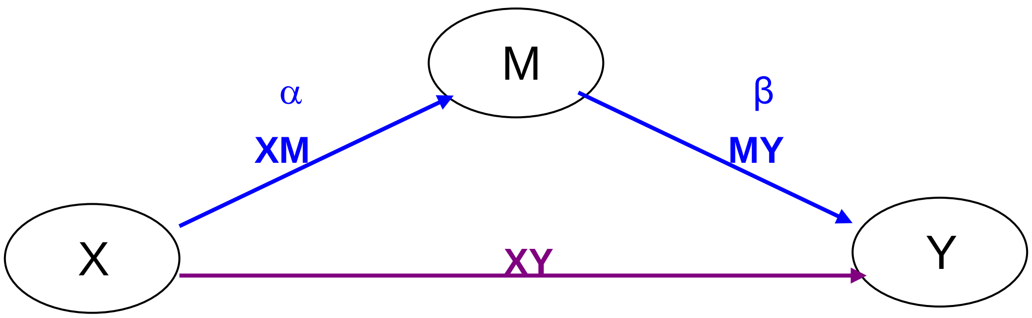
Statistical Tests of Models That Include Mediating Variables©

Consider a model that proposes that some independent variable (X) is correlated with some dependent variable (Y) not because it exerts some direct effect upon the dependent variable, but because it causes changes in an intervening or mediating variable (M), and then the mediating variable causes changes in the dependent variable. Psychologists tend to refer to the X → M → Y relationship as "mediation." Sociologists tend to speak of the "indirect effect" of X on Y through M.



MacKinnon, Lockwood, Hoffman, West, and Sheets (A comparison of methods to test mediation and other intervening variable effects, *Psychological Methods*, 2002, *7*, 83-104) reviewed 14 different methods that have been proposed for testing models that include intervening variables. They grouped these methods into three general approaches.

Causal Steps. This is the approach that has most directly descended from the work of Judd, Baron, and Kenny and which has most often been employed by psychologists. Using this approach, the criteria for establishing mediation, which are nicely summarized by David Howell (*Statistical Methods for Psychology*, 6th ed., pages 546-550) are:

- X must be correlated with Y.
- X must be correlated with M.
- M must be correlated with Y, holding constant any direct effect of X on Y.
- When the effect of M on Y is removed, X is no longer correlated with Y (complete mediation) or the correlation between X and Y is reduced (partial mediation).

Each of these four criteria are tested separately in the causal steps method:

- First you demonstrate that the zero-order correlation between X and Y (ignoring M) is significant.
- Next you demonstrate that the zero-order correlation between X and M (ignoring Y) is significant.
- Now you conduct a multiple regression analysis, predicting Y from X and M. The partial effect of M (controlling for X) must be significant.
- Finally, you look at the direct effect of X on Y. This is the Beta weight for X in the multiple regression just mentioned. For complete mediation, this Beta must be (not significantly different from) 0. For partial mediation, this Beta must be less than the zero-order correlation of X and Y.

MacKinnon et al. are rather critical of this approach. They note that it has low power. They also opine that one should not require that X be correlated with Y -- it could be that X has both a direct effect on Y and an indirect effect on Y (through M), with these two effects being equal in magnitude but opposite in sign -- in this case, mediation would exist even though X would not be correlated with Y (X would be a classical suppressor variable, in the language of multiple regression).

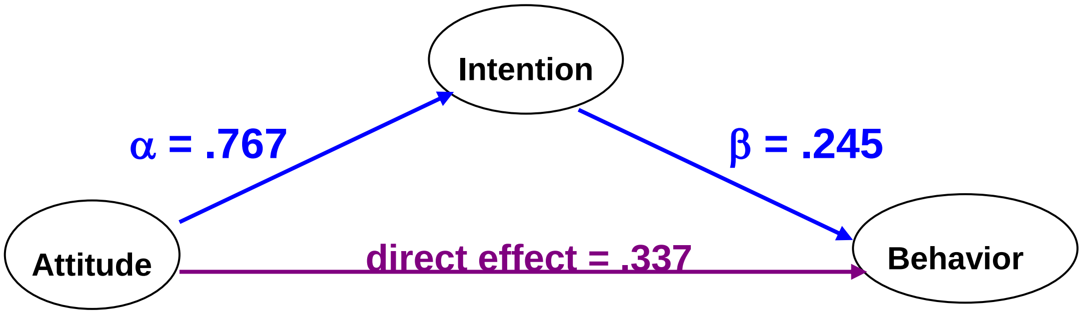
Difference in Coefficients. These methods involve comparing two regression or correlation coefficients -- that for the relationship between X and Y ignoring M and that for the relationship between X and Y after removing the effect of M on Y. MacKinnon et al. describe a variety of problems with these methods, including unreasonable assumptions and null hypotheses that can lead one to conclude that mediation is taking place even when there is absolutely no correlation between M and Y.

Product of Coefficients. One can compute a coefficient for the "indirect effect" of X on Y through M by multiplying the coefficient for path XM by the coefficient for path MY. The coefficient for path XM is the zero-order *r* between X and M. The coefficient for path MY is the Beta weight for M from the multiple regression predicting Y from X and M (alternatively one can use unstandardized coefficients).

One can test the null hypothesis that the indirect effect coefficient is zero in the population from which the sample data were randomly drawn. The test statistic (*TS*) is computed by dividing the indirect effect coefficient by its standard error, that is, $TS = \frac{\alpha\beta}{\sigma_{\alpha\beta}}$. This test statistic is usually evaluated by comparing it to the standard normal distribution. The most commonly employed standard error is **Sobel's (1982) first-order approximation**, which is computed as $\sqrt{\alpha^2\sigma_\beta^2 + \beta^2\sigma_\alpha^2}$, where α is the zero-order correlation or unstandardized regression coefficient for predicting M from X, σ_α^2 is the standard error for that coefficient, β is the standardized or unstandardized partial regression coefficient for predicting Y from M controlling for X, and σ_β^2 is the standard error for that coefficient. Since most computer programs give the standard errors for the unstandardized but not the standardized coefficients, I shall employ the unstandardized coefficients in my computations (using an interactive tool found on the Internet) below.

An alternative standard error is **Aroian's (1944) second-order exact solution**, $\sqrt{\alpha^2\sigma_\beta^2 + \beta^2\sigma_\alpha^2 + \sigma_\alpha^2\sigma_\beta^2}$. Another alternative is **Goodman's (1960) unbiased solution**, in which the rightmost addition sign becomes a subtraction sign: $\sqrt{\alpha^2\sigma_\beta^2 + \beta^2\sigma_\alpha^2 - \sigma_\alpha^2\sigma_\beta^2}$. In his text, Dave Howell employed Goodman's solution, but he made a potentially serious error -- for the MY path he employed a zero-order coefficient and standard error when he should have employed the partial coefficient and standard error.

MacKinnon et al. gave some examples of hypotheses and models that include intervening variables. One was that of Ajzen & Fishbein (1980), in which intentions are hypothesized to intervene between attitudes and behavior. I shall use here an example involving data relevant to that hypothesis. Ingram, Cope, Harju, and Wuensch (Applying to graduate school: A test of the theory of planned behavior. *Journal of Social Behavior and Personality*, 2000, *15*, 215-226) tested a model which included three "independent" variables (attitude, subjective norms, and perceived behavior control), one mediator (intention), and one "dependent" variable (behavior). I shall simplify that model here, dropping subjective norms and perceived behavioral control as independent variables. Accordingly, the mediation model (with standardized path coefficients) is:



Let us first consider the causal steps approach:

- Attitude is significantly correlated with behavior, r = .525.
- Attitude is significantly correlated with intention, r = .767.
- The partial effect of intention on behavior, holding attitude constant, falls short of statistical significance, β = .245, p = .16.
- The direct effect of attitude on behavior (removing the effect of intention) also falls short of statistical significance, β = .337, p = .056.

The causal steps approach does not, here, provide strong evidence of mediation, given lack of significance of the partial effect of intention on behavior. If sample size were greater, however, that critical effect would, of course, be statistically significant.

Now I calculate the Sobel/Aroian/Goodman tests. The statistics which I need are the following:

- The zero-order unstandardized regression coefficient for predicting the mediator (intention) from the independent variable (attitude). That coefficient = .423.
- The standard error for that coefficient = .046.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	3.390	1.519		2.231	.030
	attitude	.423	.046	.767	9.108	.000

a. Dependent Variable: intent

- The partial, unstandardized regression coefficient for predicting the dependent variable (behavior) from the mediator (intention) holding constant the independent variable (attitude). That regression coefficient = 1.065.
- The standard error for that coefficient = .751.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.075	9.056		.008	.993
	attitude	.807	.414	.337	1.950	.056
	intent	1.065	.751	.245	1.418	.162

a. Dependent Variable: behav

For Aroian's second-order exact solution,

$$TS = \frac{\alpha\beta}{\sqrt{\alpha^2\sigma_\beta^2 + \beta^2\sigma_\alpha^2 + \sigma_\alpha^2\sigma_\beta^2}} = \frac{.423(1.065)}{\sqrt{.423^2(.751)^2 + 1.065^2(.046)^2 + .046^2(.751)^2}} = 1.3935$$

What a tedious calculation that was. I just lost interest in showing you how to calculate the Sobel and the Goodman statistics by hand. Let us use Kris Preacher's dandy tool at http://quantpsy.org/sobel/sobel.htm . Just enter alpha (a), beta (b), and their standard errors and click Calculate:

	Input:		Test statistic:	*p*-value:
a	.423	Sobel test:	1.40154116	0.16105231
b	1.065	Aroian test:	1.39351588	0.16346385
s_a	.046	Goodman test:	1.40970672	0.1586263
s_b	.751	Reset all	Calculate	

Even easier (with a little bit of rounding error), just provide the *t* statistics for alpha and beta and click Calculate:

	Input:		Test statistic:	*p*-value:
t_a	9.108	Sobel test:	1.4011211	0.16117786
t_b	1.418	Aroian test:	1.39294803	0.16363551
		Goodman test:	1.40943975	0.15870518
		Reset all	Calculate	

The results indicate (for each of the error terms) a *z* of about 1.40 with a *p* of about .16. Again, our results do not provide strong support for the mediation hypothesis.

Mackinnon et al. (1998) Distribution of $\frac{\alpha\beta}{\sigma_{\alpha\beta}}$. MacKinnon et al. note one serious problem with the Sobel/Aroian/Goodman approach -- power is low due to the test statistic not really being normally distributed. MacKinnon et al. provide an alternative approach. They used Monte Carlo simulations to obtain critical values for the test statistic. A table of these critical values for the test statistic which uses the Aroian error term (second order exact formula) is available at http://www.public.asu.edu/~davidpm/ripl/freqdist.second.pdf . This table is also available from Dr. Karl Wuensch. Please note that the table includes sampling distributions both for populations where the null hypothesis is true (there is no mediating effect) and where there is a mediating effect. Be sure you use the appropriate (no mediating effect) portion of the table to get a *p* value from your computed value of the test statistic. The first four pages of the table give the percentiles from 1 to

100 under the null hypothesis when all variables are continuous. Later in the table (pages 17-20) is the same information for the simulations where the independent variable was dichotomous and there was no mediation.

When using the Aroian error term, the .05 critical value is ± 0.9 -- that is, if the absolute value of the test statistic is .9 or more, then the mediation effect is significant. Using the Sobel error term, the .05 critical value is ± 0.97. MacKinnon et al. refer to the test statistic here as z′, to distinguish it from that for which one (inappropriately) uses the standard normal PDF to get a *p* value. With the revised test, we do have evidence of a significant mediation effect.

The table can be confusing. Suppose you are evaluating the product of coefficients test statistic computed with Aroian's second-order exact solution and that your sample size is approximately 200. For the traditional two-tailed .05 test, the critical value of the test statistic is that value which marks off the lower 2.5% of the sampling distribution and the upper 2.5% of the sampling distribution. As noted at the top of the table, the absolute critical value is approximately .90. The table shows that the 2nd percentile has a value of -.969 and the 3rd a value of -.871. The 2.5th percentile is not tabled, but as noted earlier, it is approximately -.90. The table shows that the 97th percentile has a value of .868 and the 98th a value of .958. The 97.5th percentile is not tabled, but, again, it is approximately .90. So, you can just use .90 as the absolute critical value for a two-tailed .05 test. If you want to report an "exact" *p* (which I recommend), use the table to find the proportion of the area under the curve beyond the your obtained value of the test statistic and then, for a two-tailed test, double that proportion. For example, suppose that you obtained a value of 0.78. From the table you can see that this falls near the 96th percentile -- that is, the upper-tailed *p* is about .04. For a two-tailed *p*, you double .04 and then cry because your *p* of .08 is on the wrong side of that silly criterion of .05.

Mackinnon et al. (1998) Distribution of Products. With this approach, one starts by converting both critical paths (α and β in the figure above) into *z* scores by dividing their unstandardized regression coefficients by the standard errors (these are, in fact, the *t* scores reported in typical computer output for testing those paths). For our data, that yields $Z_\alpha Z_\beta = 9.108 * 1.418 = 12.915$. For a .05 nondirectional test, the critical value for this test statistic is 2.18. Again, our evidence of mediation is significant.

MacKinnon et al. used Monte Carlo techniques to compare the 14 different methods' statistical performance. A good method is one with high power and which keeps the probability of a Type I error near its nominal value. They concluded that the best method was the Mackinnon et al. (1998) distribution of $\dfrac{\alpha\beta}{\sigma_{\alpha\beta}}$ method, and the next best method was the Mackinnon et al. (1998) distribution of products method.

Bootstrap Analysis. Patrick Shrout and Niall Bolger published an article, "Mediation in Experimental and Nonexperimental Studies: New Procedures and Recommendations," in the *Psychological Bulletin* (2002, *7*, 422-445), in which they recommend that one use bootstrap methods to obtain better power, especially when sample sizes are not large. They included, in an appendix (B), instructions on how to use EQS or AMOS to implement the bootstrap analysis. Please note that Appendix A was corrupted when printed in the journal. A corrected appendix can be found at http://www.psych.nyu.edu/couples/PM2002.

Kris Preacher and Andrew Hayes have provided SAS and SPSS macros for bootstrapped mediation analysis, and recommend their use when you have the raw data, especially when sample size is not large. Here I shall illustrate the use of their SPSS macro.

- I bring the raw data from Ingram's research into SPSS.

- I bring the sobel_SPSS.sps syntax file into SPSS.
- I click Run, All.
- I enter into another syntax window the command
 "SOBEL y=behav / x=attitude / m=intent /boot=10000."
- I run that command. Now SPSS is using about 50% of the CPU on my computer and I hear the fan accelerate to cool down its guts. Four minutes later the output appears:

```
Run MATRIX procedure:

DIRECT AND TOTAL EFFECTS
            Coeff      s.e.         t  Sig(two)
b(YX)      1.2566     .2677    4.6948     .0000
b(MX)       .4225     .0464    9.1078     .0000
b(YM.X)    1.0650     .7511    1.4179     .1617
b(YX.M)     .8066     .4137    1.9500     .0561

INDIRECT EFFECT AND SIGNIFICANCE USING NORMAL DISTRIBUTION
            Value      s.e.  LL 95 CI  UL 95 CI         Z  Sig(two)
  Sobel     .4500     .3231   -.1832    1.0831    1.3929     .1637

BOOTSTRAP RESULTS FOR INDIRECT EFFECT
             Mean      s.e.  LL 95 CI  UL 95 CI  LL 99 CI  UL 99 CI
 Effect     .4532     .2911   -.1042    1.0525   -.2952    1.2963

SAMPLE SIZE
       60

NUMBER OF BOOTSTRAP RESAMPLES
    10000

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

If you look at the bootstrapped confidence (95%) interval for the indirect effect (in unstandardized units), -0.1042 to 1.0525, you see that bootstrap tells us that the indirect effect is not significantly different from zero.

X and Y Temporally Distant. Shrout and Bolger also discussed the question of whether or not one should first verify that there is a zero-order correlation between X and Y (Step 1 of the causal steps method). They argued that it is a good idea to test the zero-order correlation between X and Y when they are proximal (temporally close to one another), but not when they are distal (widely separated in time). When X and Y are distal, it becomes more likely that the effect of X on Y is transmitted through additional mediating links in a causal chain and that Y is more influenced by extraneous variables. In this case, a test of the mediated effect of X on Y through M can be more powerful than a test of the zero-order effect of X on Y.

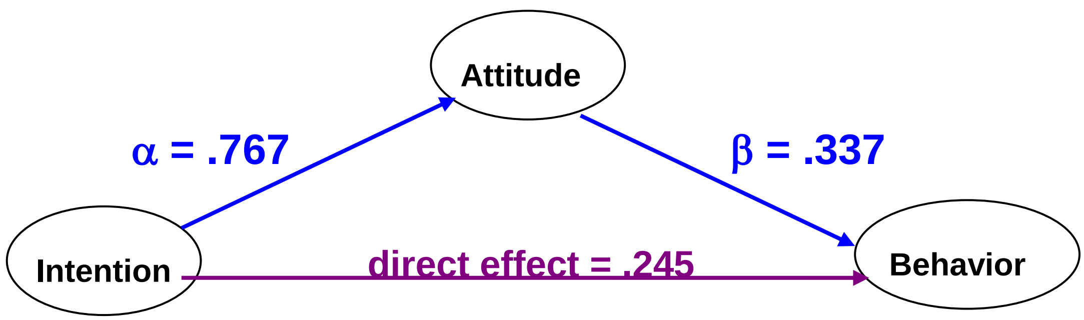
Shrout and Bolger also noted that the direct effect of X on Y could be opposite in direction of the indirect effect of X on Y, leading to diminution of the zero-order correlation between X and Y. Requiring as a first step a significant zero-order correlation between X and Y is not recommended when such suppressor effects are considered possible. Shrout and Bolger gave the following example:

- X is the occurrence of an environmental stressor, such as a major flood, and which has a direct effect of increasing Y
- Y is the stress experienced by victims of the flood.
- M is coping behavior on part of the victim, which is initiated by X and which reduces Y.

Partial Mediation. Shrout and Bolger also discussed three ways in which one may obtain data that suggest partial rather than complete mediation:

- X may really have a direct effect upon Y in addition to its indirect effect on Y through M.
- X may have no direct effect on Y, but may have indirect effects on Y through M_1 and M_2. If, however, M_2 is not included in the model, then the indirect effect of X on Y through M_2 will be mistaken as being a direct effect of X on Y.
- There may be two subsets of subjects. In the one subset there may be only a direct effect of X on Y, and in the second subset there may be only an indirect effect of X on Y through M.

Causal Inferences from Nonexperimental Data? Please note that no variables were manipulated in the research by Ingram et al. One might reasonably object that one can never establish with confidence the veracity of a causal model by employing nonexperimental methods. With nonexperimental data, the best we can do is to see if our data fit well with a particular causal model. We must keep in mind that other causal models might fit the data equally well or even better. For example, we could consider the following alternative model, which fits the data equally well:

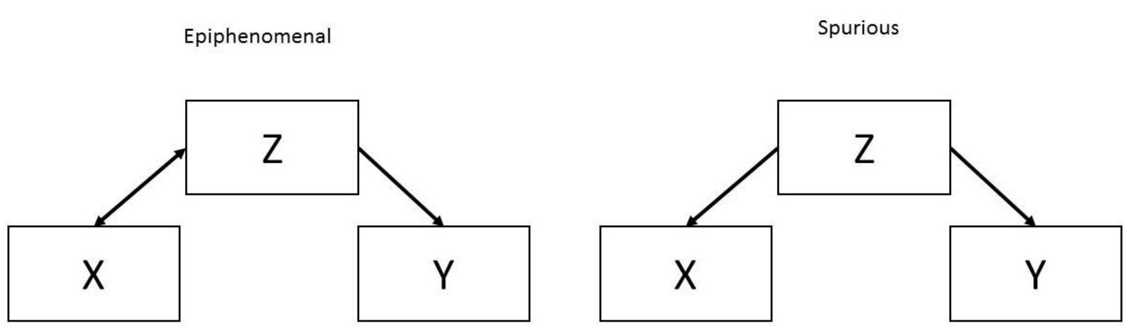


When X is not experimentally manipulated, then things get even worse. Absent random assignment to values of X, all of the associations in a mediation model are susceptible to confounding and epiphenomenal association ... (Andrew F. Hayes, p. 122, Introduction to Mediation, Moderation, and Conditional Process Analysis, 2nd edition, 2018)

An association between X and Y is epiphenomenal if X is correlated with a cause of Y but not itself causally influence Y. (p. 46).

Two variables are spuriously correlated if their association is induced as a result of a shared cause. (p. 46)

If X is not determined through manipulation and random assignment, then any sequence of causal ordering of X, M, and Y must be entertained as a potential candidate for the direction of causal flow. (p. 130)



If X causes Z, then Z is a mediator, but since Z is not in our model, its effect goes into the direct effect.

If Z causes X, we have a classic "third variable problem," and the XY correlation is spurious.

If we cannot decide whether X causes Z or Z causes X, we have an epiphenomenal association.

Models containing intervening variables, with both direct and indirect effects possible, can be much more complex than the simple trivariate models discussed here. Path analysis is one technique for studying such complex causal models. For information on this topic, see my document "An Introduction to Path Analysis."

Direct, Indirect, and Total Effects

It is always a good idea to report the (standardized) direct, indirect, and total effects, all of which can be obtained from the path coefficients. Using our original model, the direct effect of attitude is .337. The indirect effect is (.767)(.245) = .188. The total effect is the simply the sum of the direct and indirect effect, .337 + .188 = .525. The zero-order correlation between attitude and behavior was .525, so what we have done here is to partition that correlation into two distinct parts, the direct effect and the indirect effect.

Example of How to Present the Results

Have a look at the articles which I have made available in BlackBoard. While they are more complex than the analysis presented above, they should help you learn how mediation analyses are presented in the literature. Below is an example of how a simple mediation analysis can be presented.

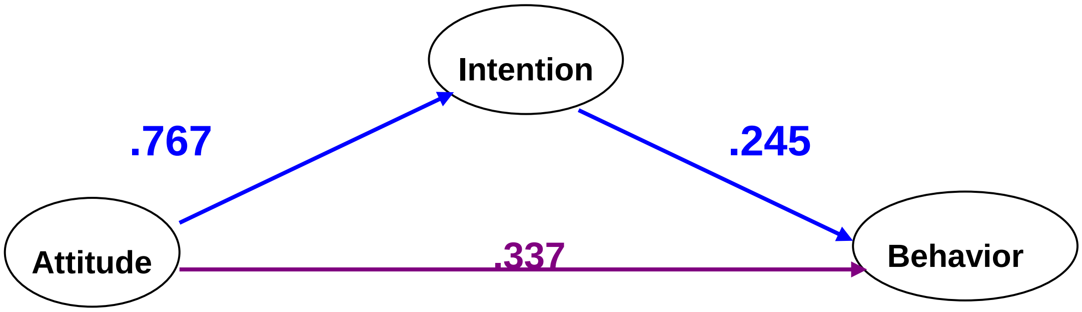
Example Presentation of Results of a Simple Mediation Analysis

As expected, attitude was significantly correlated with behavior, r = .525, p < .001., 95% CI [.31, .69]. Sequential correlation analysis was employed to investigate the involvement of intention as a possible mediator of the relationship between attitude and behavior.

Attitude was found to be significantly related to the intention, r = .767, p < .001, 95% CI [.64, .86]. Behavior was significantly related to a linear combination of attitude and intention, $F(2, 57)$ = 12.22, p < .001, R = .548, 95% CI [.32, .69]. Neither attitude (β = .337, p = .056) nor intention (β = .245, p = .162) had a significant partial effect on behavior.

Aroian's test of mediation indicated that intention significantly mediated the relationship between the attitude and behavior, TS = 1.394, p < .02. It should be noted that the p value was not obtained from the standard normal distribution but rather from the table provided by MacKinnon, Lockwood, Hoffman, West, & Sheets (2002), available at http://www.public.asu.edu/~davidpm/ripl/freqdist.pdf. [Note: It would be better practice to report the bootstrapping test of the indirect effect].

The mediation model is illustrated below. The indirect effect of attitude on behavior, (.767)(.245) = .188, and it direct effect is .337, yielding a total effect coefficient of .525 (not coincidentally equal to the zero-order correlation between attitude and behavior). Accordingly, .188/.525, 35.8% of the effect of attitude on behavior is mediated through intention, and .337/.525 = 62.4% is direct. Of course, this direct effect may include the effects of mediators not included in the model. [Note: There are some potentially serious problems with the ratio of the indirect effect to the total effect – see Hayes (2013), pages 188-193. Probably best to stick with Preacher & Kelley's κ^2].



I am not terribly fond of the practice of computing percentages for the direct and indirect effects, but many researchers do report such percentages. I don't know how they would handle a case where the indirect and direct effects are in different directions. For example, if the IE were +0.20 and the DE were -0.15, the total effect would be +0.05. Would the IE as a percentage be 400% ? Would the DE as a percentage be -300% ?

Process Hayes for Mediation Analysis

Andy Hayes has made available SAS and SPSS code that makes it easy to do mediation analysis. Give it a try.

1. If you have not already done so, go to the Process Page and download the Process zip file. In the zip are, among other items, process.sas and process.sps. Bring the process.sas program into SAS and run it. It is a macro, so once you run it you can go ahead and close the program (don't close SAS yet), if you wish.

2. Go to my Stat Data page and download the Path-Ingram.xlsx file. Import it into SAS. Give the data the SAS name "Ingram." Then enter the below code into the SAS editor and run it.

```
%process (data=Ingram,vars=Behavior Attitude
Intent,y=Behavior,x=Attitude,m=Intent,model=4,
total=1, boot=10000, effsize=1);
```

- Data= points to the SAS data file which you just imported
- Vars= identifies the variables
- Y= identifies the outcome variable
- X= identifies the focal predictor variable
- M= identifies the mediator variable
- Model=1 identifies the simple moderation model – see the templates document
- Total=1 asks for total, direct, and indirect effects
- Boot=10,000 asks SAS to use 10,000 bootstrap samples
- Effsize=1 asks SAS to compute several effect size statistics

Here is the output, with annotations:

The SAS System

************************* PROCESS Procedure for SAS Release 2.10 ************************

Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

Documentation available in Hayes (2013). www.guilford.com/p/hayes3

****

Model and Variables	
Model =	4
Y =	BEHAVIOR
X =	ATTITUDE
M =	INTENT

Sample size:
60

****

Outcome: INTENT

Model Summary					
R	**R-sq**	**F**	**df1**	**df2**	**p**
0.7671	0.5885	82.9524	1.0000	58.0000	0.0000

Model						
	coeff	**se**	**t**	**p**	**LLCI**	**ULCI**
Constant	3.3895	1.5193	2.2309	0.0296	0.3483	6.4308
ATTITUDE	0.4225	0.0464	9.1078	0.0000	0.3296	0.5154

This is the unstandardized path coefficient for predicting intention from attitude. It is significant. Hayes prefers unstandardized coefficients. I prefer standardized coefficients. To get standardized coefficients, simply standardize your variables prior to analysis (use Proc Standard).

****

Outcome: BEHAVIOR

Model Summary					
R	**R-sq**	**F**	**df1**	**df2**	**p**
0.5478	0.3001	12.2177	2.0000	57.0000	0.0000

Model						
	coeff	**se**	**t**	**p**	**LLCI**	**ULCI**
constant	0.0747	9.0562	0.0082	0.9934	-18.0600	18.2095
INTENT	1.0650	0.7511	1.4179	0.1617	-0.4391	2.5691
ATTITUDE	0.8066	0.4137	1.9500	0.0561	-0.0217	1.6350

Here we have the other two unstandardized path coefficients. Neither is significant.

************************** TOTAL EFFECT MODEL **************************

Outcome: BEHAVIOR

Model Summary					
R	**R-sq**	**F**	**df1**	**df2**	**p**
0.5248	0.2754	22.0410	1.0000	58.0000	0.0000

Model						
	coeff	**se**	**t**	**p**	**LLCI**	**ULCI**
constant	3.6845	8.7663	0.4203	0.6758	-13.8632	21.2323
ATTITUDE	1.2566	0.2677	4.6948	0.0000	0.7208	1.7924

This is the unstandardized total effect.

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

Total effect of X on Y					
Effect	**SE**	**t**	**p**	**LLCI**	**ULCI**
1.2566	0.2677	4.6948	0.0000	0.7208	1.7924

Direct effect of X on Y					
Effect	**SE**	**t**	**p**	**LLCI**	**ULCI**
0.8066	0.4137	1.9500	0.0561	-0.0217	1.6350

We saw this earlier, in the "Outcome: Behavior" model.

Indirect effect of X on Y				
	Effect	**Boot SE**	**BootLLCI**	**BootULCI**
INTENT	0.4500	0.2977	-0.1081	1.0815

To get the unstandardized indirect effect, we simply multiply the two coefficients for the path leading from Attitude to Behavior -- .4225(1.0650) = .45. Notice that the confidence interval includes zero, so it is not significant.

Partially standardized indirect effect of X on Y				
	Effect	Boot SE	BootLLCI	BootULCI
INTENT	0.0270	0.0180	-0.0079	0.0639

Here Y has been standardized, but X and M have not. If X is dichotomous, this statistic tells you by how many standard deviations the two groups differ because of the indirect effect of X on Y.

Completely standardized indirect effect of X on Y				
	Effect	Boot SE	BootLLCI	BootULCI
INTENT	0.1879	0.1263	-0.0394	0.4667

This is the standardized indirect effect coefficient, which we previous calculated as the product of the two standardized path coefficients.

Ratio of indirect to total effect of X on Y				
	Effect	Boot SE	BootLLCI	BootULCI
INTENT	0.3581	0.2720	-0.0838	0.9975

To compute this statistic by hand, simply divide the indirect effect by the total effect -- .45/1.2566 = .358. Although this sounds like the proportion of the total effect that is indirect, this statistic is not really a proportion, as it can exceed one or be negative.

Ratio of indirect to direct effect of X on Y				
	Effect	Boot SE	BootLLCI	BootULCI
INTENT	0.5578	56.8327	-0.1420	9.2604

To compute this statistic by hand, simply divide the indirect effect by the direct effect -- .45/.8806 = .558. Although both of these ratio measures have been quite popular, they have a number of limitations, including being unstable across samples unless the sample size is very large. See Preacher and Kelley (2011) for more details.

R-squared mediation effect size				
	Effect	Boot SE	BootLLCI	BootULCI
INTENT	0.2287	0.0744	0.0941	0.3845

This is "the variance in Y that is common to both X and M but that can be attributed to neither alone." One limitation of this statistic is that it can be negative in the presence of suppressor effects. Again, see Preacher and Kelley (2011).

Preacher and Kelley (2011) Kappa-squared				
	Effect	Boot SE	BootLLCI	BootULCI
INTENT	0.1416	0.0813	0.0094	0.3163

This statistic is the ratio of the indirect effect to the maximum value that the indirect effect could assume given the constraints imposed by variances and covariance of X, M, and Y. It has the advantage of ranging from 0 to 1, as a proportion should. Notice that for our data this statistic is significantly greater than zero. Work by Wen and Fan (2015) has cast doubt on the validity of this statistic. Preacher and Kelley made an error in the computation of the maximum value of the indirect effect.

Wen and Fan (2015) concluded "The main motivation for developing an effect size measure is to overcome the difficulty introduced by the arbitrariness of scales of the variables in an analysis (Fan & Konold, 2010), so that (a) we can meaningfully interpret the observed effect, and (b) findings from different studies could be meaningfully synthesized (e.g., meta-analysis). The indirect effect ab, when in standardized form, is already on a standardized scale that is directly interpretable. For example, if the total effect c = 0.5, indirect effect ab = 0.2, and direct effect c= = 0.3 (all in standardized form), we can interpret these as: corresponding to a change of one standard deviation in X, Y is expected to change by 0.5 standard deviation, of which 0.2 would be the indirect effect of X on Y via the mediator M, and the remaining 0.3 is the direct effect of X on Y. Because the standardized form of the mediation effect ab is already interpretable, and it could be used for synthesis across studies, it appears that there is no fundamental motivation for developing another scale-free index to represent ab, because ab in standardized form is already scale-free.

A Big Fly in the Ointment

I have always been uncomfortable with tests of causal models that do not involve the manipulation of any variables and employ data that are all gathered at the same time. If X causes M and then M causes Y, should not X, M, and Y be measured at different times, allowing time for X to cause changes in M and M to cause changes in Y? Maxwell and Cole (2007) argue that mediation models should be tested only with longitudinal data. Figure 3 from their article, below, illustrates a longitudinal mediation model.

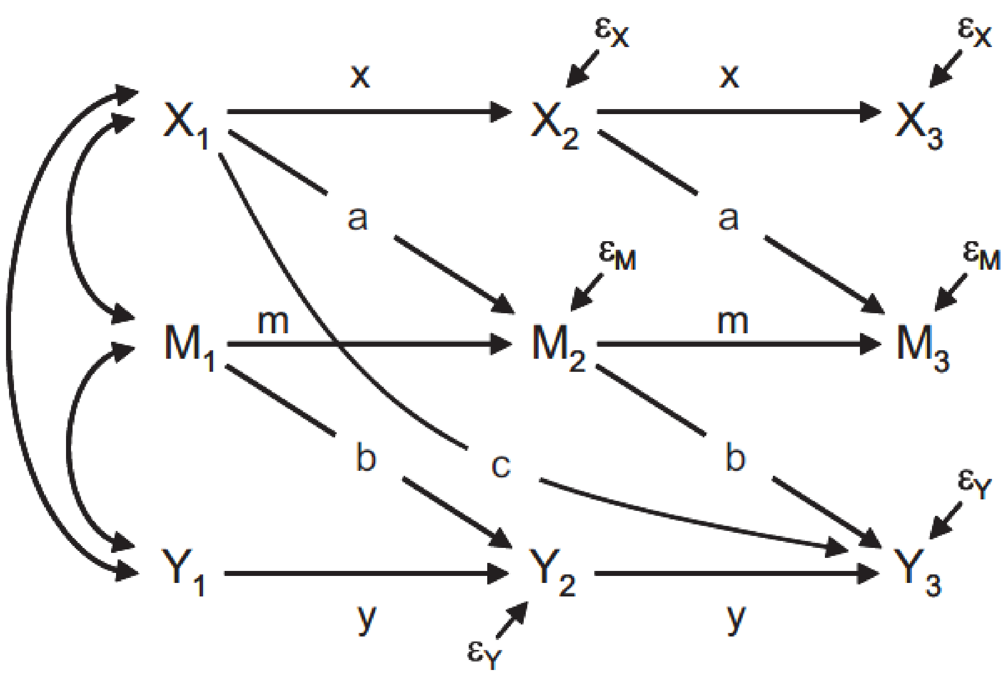


Figure 3. Longitudinal mediation model. X = independent variable; M = mediator; Y = dependent variable.

Paths *x*, *m*, and *y* represent the "autoregressive" effects of a variable at one time on the same variable at a later time. Paths *a*, *b*, and *c* represent the direct effects of one variable at one time on another variable later in time.

Maxwell and Cole (2007) and Maxwell, Cole, and Mitchell (2011) have demonstrated that mediation analysis of cross-sectional data (data gathered all at one time) typically produces (very) biased estimates of the effects of mediators. They recommend that mediation hypotheses be tested only with longitudinal data. They also show that most of the research articles involving tests of mediation models have not used fully longitudinal data. Partially longitudinal data are when X and M or M and Y, but not both X and M and M and Y, are separated in time.

References

- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford.
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*, 2nd ed. New York, NY: Guilford.
- Howell, D. C. (2013). ***Statistical methods for psychology*** (8th ed.). Belmont, CA: Cengage Wadsworth.
- MacKinnon, D. P., Lockwood, C., & Hoffman, J. (1998, June). *A new method to test for mediation*. Paper presented at the annual meeting of the Society for Prevention Research, Park City, UT.
- MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M,, West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83-104.
- Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, *12*, 23-44.
- Maxwell, S. E., Cole, D. A., & Mitchell, M. A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model, *Multivariate Behavioral Research, 46*, 816-841.
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, *16*, 93-115.
- Shrout, P. E., & Bolger, N. (2002) Mediation in experimental and nonexperimental studies: New procedures and recommendations, *Psychological Bulletin*, *7*, 422-445.
- Wen, Z., & Fan, X. (2015). Monotonicity of effect sizes: Questioning kappa-squared as mediation effect size measure. *Psychological Methods*, *20*, 193-203.

Links

- Internet Resources
- Return to Wuensch's Statistics Lessons Page