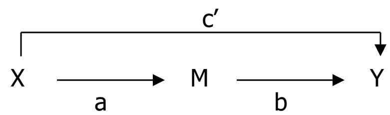
Testing Mediation with Regression Analysis

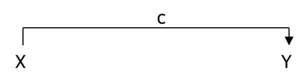
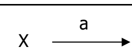
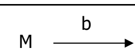
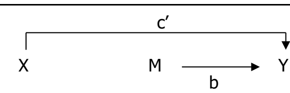
Mediation is a hypothesized causal chain in which one variable affects a second variable that, in turn, affects a third variable. The intervening variable, M, is the mediator. It "mediates" the relationship between a predictor, X, and an outcome. Graphically, mediation can be depicted in the following way:

Paths *a* and *b* are called direct effects. The mediational effect, in which X leads to Y through M, is called the *indirect effect*. The indirect effect represents the portion of the relationship between X and Y that is mediated by M.

Testing for mediation

Baron and Kenny (1986) proposed a four-step approach in which several regression analyses are conducted and significance of the coefficients is examined at each step. Take a look at the diagram below to follow the description (note that *c′* could also be called a direct effect).



	Analysis	*Visual Depiction*
Step 1	Conduct a simple regression analysis with X predicting Y to test for path *c* alone, $Y = B_0 + B_1X + e$	
Step 2	Conduct a simple regression analysis with X predicting M to test for path *a*, $M = B_0 + B_1X + e$.	
Step 3	Conduct a simple regression analysis with M predicting Y to test the significance of path *b* alone, $Y = B_0 + B_1M + e$.	
Step 4	Conduct a multiple regression analysis with X and M predicting Y, $Y = B_0 + B_1X + B_2M + e$	

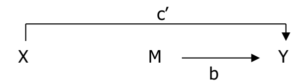
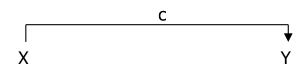
The purpose of Steps 1-3 is to establish that zero-order relationships among the variables exist. If one or more of these relationships are nonsignificant, researchers usually conclude that mediation is not possible or likely (although this is not always true; see MacKinnon, Fairchild, & Fritz, 2007). Assuming there are significant relationships from Steps 1 through 3, one proceeds to Step 4. In the Step 4 model, some form of mediation is supported if the effect of M (path *b*) remains significant after controlling for X. If X is no longer significant when M is controlled, the finding supports *full mediation*. If X is still significant (i.e., both X and M both significantly predict Y), the finding supports *partial mediation*.

Calculating the indirect effect

The above four-step approach is the general approach that most researchers used to follow traditionally. There are potential problems with this approach, however. One problem is that when the ab effect is off opposite sign to the c effect, known as inconsistent mediation (also a suppression effect), the four-step approach may miss a true mediational effect (Mackinnon et al., 2000). A second problem is that, in some instances, it is possible that the overall X -> Y effect is not significant a test of the indirect effect may be significant (O'Rourke & MacKinnon, 2015). The standard error for the overall (c) effect tends to be larger than the standard error for the indirect path (ab). This means that the Barron and Kenny (and Judd & Kenny) stepped approach tends to miss some true mediation effects (Type II errors; MacKinnon et al., 2007; O'Rourke & MacKinnon, 2015). **An alternative, and**

preferable approach, is to calculate the indirect effect and test it for significance using confidence intervals (Mackinnon et al., 2023).

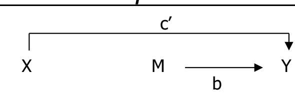
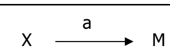
The regression coefficient for the indirect effect represents the change in Y for every unit change in X that is mediated by M. There are two ways to estimate the indirect coefficient. Judd and Kenny (1981) suggested computing the difference between two regression coefficients. To do this, two regressions are required.

Judd & Kenny Difference of Coefficients Approach		
	Analysis	*Visual Depiction*
Model 1	$Y = B_0 + B_1X + B_2M + e$	
Model 2	$Y = B_0 + BX + e$	

The approach involves subtracting the partial regression coefficient obtained in Model 1, B_1 from the simple regression coefficient obtained from Model 2, B. Note that both represent the effect of X on Y but that B is the coefficient from the simple regression and B_1 is the partial regression coefficient from a multiple regression. The indirect effect is the difference between these two coefficients:

$$B_{indirect} = B - B_1 .$$

An equivalent approach calculates the indirect effect by multiplying two regression coefficients (Sobel, 1982). The two coefficients are obtained from two regression models.

Sobel Product of Coefficients Approach		
	Analysis	*Visual Depiction*
Model 1	$Y = B_0 + B_1X + B_2M + e$	
Model 2	$M = B_0 + BX + e$	

Notice that Model 2 is a different model from the one used in the difference approach. In the Sobel approach, Model 2 involves the relationship between X and M. A product is formed by multiplying two coefficients together, the partial regression effect for M predicting Y, B_2, and the simple coefficient for X predicting M, B:

$$B_{indirect} = (B_2)(B)$$

As it turns out, the Kenny and Judd difference of coefficients approach and the Sobel product of coefficients approach yield identical values for the indirect effect (MacKinnon, Warsi, & Dwyer, 1995). Note: regardless of the approach you use (i.e., difference or product) be sure to use *unstandardized* coefficients if you do the computations yourself.

Statistical tests of the indirect effect

Once the regression coefficient for the indirect effect is calculated, it needs to be tested for significance or a confidence interval needs to be constructed. There has been considerable controversy about the best way to estimate the standard error used in the significance test or confidence interval, however, and there are quite a few proposed approaches to calculation of standard errors. One of the reasons for needing alternative methods is that the sampling distribution of the indirect effect may not be normal (Kisbu-Sakarya, MacKinnon, and Miočević, 2014), and this has led to more emphasis on confidence intervals, which can be constructed to be asymmetric.

There are two general approaches to testing significance of the indirect effect that appear to perform better than the alternatives in simulation studies—bootstrap methods (sometimes called "nonparametric resampling") and the Monte Carlo method (sometimes called "parametric resampling" or "numerical integration" method). For the bootstrap method, software for testing indirect effects generally offers two options. One, referred to as "percentile" bootstrap, involves confidence intervals using usual sampling distribution cutoffs without explicit bias corrections. The other is "bias-corrected" (or sometimes BC). The bias-corrected bootstrap estimates correct for a bias in the indirect coefficient using the average estimate from the bootstrap samples. In addition to just correcting the indirect coefficient, an option may be to use confidence limits with a graded correction in the standard deviation across potential values of the indirect coefficient, referred to as an "accelerated" bias-corrected bootstrap. The Monte Carlo approach is another approach (not widely available currently, however) is a bit different from the bootstrap approaches. The Monte Carlo approach involves computation of the indirect effect and the standard error estimates for the separate coefficients for the full sample. Resampling is then used to estimate the standard errors for the indirect effects using these values. The bias-corrected bootstrap method may result in Type I error rates that are slightly higher than the percentile bootstrap method (Biesanz, Falk, & Savalei, 2010; Fritz, Taylor, & MacKinnon, 2012). Tofighi and MacKinnon (2016) find that both the **percentile bootstrap confidence intervals** and the **Monte Carlo method provide good tests** with good Type I error rates and statistical power but that the Monte Carlo approach had somewhat better power in one circumstance. Standardized coefficients can be computed, using the products of standardized coefficients from Model 1 and Model 2 above, $(\beta_2)(\beta)$, though they may not be reported by the software program, and other methods such as the ratio of indirect to total effect can also be computed (see Preacher & Kelley, 2011, for a review). You may also hear about an approach derived from the stepped approach proposed by Baron and Kenny (1986), referred to as joint-significance testing approach, in which the a and b effect are tested simultaneously and the indirect effect is considered significant if neither confidence interval contains zero (Biesanz, Falk, & Savalei, 2010; Fritz & MacKinnon, 2007). This approach also provides a good Type I error and power balance (MacKinnon et al., 2002).

Recommended approach. The **indirect effect should be tested** and researchers should not stop because the direct effect of X on Y is nonsignificant if there is a mediational hypothesis (O'Rourke & Mackinnon, 2018). When testing the indirect effect, **use bootstrap or Monte Carlo confidence intervals**, which can be asymmetric, instead of using the significance test. The significance test of $b_{indirect} / SE_{b_{indirect}}$, even in conjunction with a resampling (e.g., boostrap) process assumes a symmetric sampling distribution.

For bootstrap tests, use the percentile bootstrap. The bias-corrected bootstrap method may result in Type I error rates that are higher than the percentile bootstrap method (Biesanz, Falk, & Savalei, 2010; Chen & Fritz, 2021; Falk & Biesanz, 2015; Fritz, Taylor, & MacKinnon, 2012; Tofighi and MacKinnon, 2016; Valente, Gonzalez, Miočević, & MacKinnon, 2016). Type I error for the bias-corrected approach is higher when sample size is larger, the paths are smaller, either the a or b path is close to zero. These studies also generally show that that both the percentile bootstrap confidence intervals and the Monte Carlo method provide tests with good Type I error rates under the most conditions. Note that it is important to use the confidence intervals rather than a significance test constructed from the parameter estimated divided by the standard error because of the asymmetry of the sampling distribution. The bias-corrected bootstrap method is often shown to have higher statistical power (and sometimes the joint-significance test; Yzerbyt et al., 2018), but this is at the cost of inflated Type I error (Valente et al., 2016). Tofighi and MacKinnon (2016) show that Monte Carlo approach had somewhat better power for an indirect effect with several sequential mediators (i.e., a product of four paths). Although less often included in comparison of approaches, Yzerbyt and colleagues (2018) showed that the joint-significance test controlled Type I errors well and had comparable statistical power to the percentile bootstrap and Monte Carlo methods. The joint-significance approach assumes the a and b effect are independent, and this may lead to less accuracy of this approach when the two effects are correlated. Overall, my read is that the balance of power and Type I error across more circumstances, combined its wide availability and ease of use, leads to a recommendation to use the percentile bootstrap confidence interval or the Monte Carlo method as a usual approach.

Effect size measures. Standardized coefficients can be computed (Lachowicz, Preacher, & Kelley, 2018; Miočević, O'Rourke, MacKinnon, & Brown (2018), but they are often not computed with macros or packages that test indirect effects. Computation is simple by hand using the indirect effect (ab) and the ratio of standard deviations of X (predictor) and Y (final outcome) in the usual standardized coefficient equation (Miočević, et al., 2018), $\beta_{indirect} = B_{indirect}\left(sd_x / sd_y\right)$. Another simple method is to use the products of standardized coefficients from Model 1 and Model 2 above, (β_2)(β). Or one can pre-standardize the variables and run the analysis, although you should ignore the significance tests if using this approach. Other methods such as the ratio of indirect to total effect have been suggested for gauging the magnitude of effect (see Preacher & Kelley, 2011, for a review), although MacKinnon and colleagues (MacKinnon, Warsi, & Dwyer, 1995) found this measure to be unstable with smaller sample sizes (e.g., < 400). Lachowicz and colleagues (2018) propose a measure of effect size (upsilon, υ) related to the standardized indirect effect but which can be used with binary mediators and power analysis.

Power. Statistical tests of indirect effects often suffer from low power (MacKinnon et al., 2002; MacKinnon, Lockwood, & Williams, 2004), so researchers should plan for larger sample sizes. There are a couple of reasons for lower power of indirect effects compared with direct effects. One reason is that, because the indirect effect is product of two regression coefficients, coefficient a (X predicting M) and coefficient b (M predicting Y), the effect size and power of the indirect effects also are functions of the product the direct effects (Fritz & MacKinnon, 2007; Wang & Xue, 2021). Smaller effects sizes may stem from a variety of other factors as well (Walters, 2019), such as measurement error which may be compounded because there are three or more variables involved in the indirect pathway. The nonnormal sampling distribution for indirect coefficients also impacts power to determine significance. Power for the indirect effect depends the a and b effect in some complex ways. For small b effects, moderately-sized a coefficients, more than smaller- and larger-sized a coefficients, may lead to a stagnation of power, in which power is worse than expected given the size of the coefficient (Fritz, Taylor, & MacKinnon, 2012; Kenny & Judd, 2014).

Software. There are several possible computer methods of estimating and testing indirect effects, and I will focus on two (the PROCESS macro for SPSS and the `mediation` package for R) that use the percentile bootstrap method in the subsequent handout "Testing Mediation with Regression Analysis Example." Both of these methods estimate the indirect tests and confidence limits from the model and data in a single step.[1] The `RMediation` package (Tofighi & MacKinnon, 2011) will estimate confidence limits with the Monte Carlo method.

---

[1] Of course, the indirect coefficient can be computed by hand (as described above) from the product of B_2 (from Model 1) and B (from Model 2). The standard error can also be computed by hand and a significance test or confidence limits can be obtained in the usual manner. The standard errors can be computed more conveniently using Preacher and Leonardelli's online calculator, http://quantpsy.org/sobel/sobel.htm, if the appropriate (direct effect) unstandardized coefficients and their standard errors are entered. The online calculator returns Sobel, Goodman, and Aroian versions of the indirect standard error. Although the Sobel and Aroian tests will likely give fairly accurate statistical tests and confidence limits for large sample sizes, the bootstrap or Monte Carlo tests are preferred for best accuracy in the widest range of circumstances.

Many structural equation modeling packages, such as Mplus or R lavaan, can conduct the same types of tests of the indirect effects. SEM tests the paths specified in the model, and, upon request, can estimate confidence intervals for any indirect effects using bootstrap methods. The percentile bootstrap can be specified in Mplus using `cinterval (bootstrap)` and in lavaan using `ci = TRUE, boot.ci.type = "norm", level = 0.95`. The `monteCarloMed` function in the `semTools` R package will test the indirect effect with the Monte Carlo approach. For measured variables and continuous variables, this approach is equivalent to the regression approach. But SEM makes it possible to test more complicated models, with multiple mediators or multiple links in the chain, or latent variables, all tested as part of the usual model testing process rather than use of regressions conducted in separate steps. In addition, the SEM analysis approach provides model fit information that provides information about consistency of the hypothesized mediational model to the data (more on this issue later). Measurement error is a potential concern in mediation testing because of attenuation of relationships (Baron & Kenny, 1986; Fritz, Kenny, & MacKinnon, 2016; Gonzalez et al, 2021; VanderWeele, Valeri, & Ogburn, 2012), and the SEM approach can address this problem by removing measurement error from the estimation of the relationships among the variables when latent variables are incorporated

Online resources

David Kenny also has a webpage on mediation: http://davidakenny.net/cm/mediate.htm
Preacher's Sobel test calculator: http://quantpsy.org/sobel/sobel.htm
Hayes's PROCESS macro: https://processmacro.org/index.html
Mediation package in R: https://cran.r-project.org/web/packages/mediation/mediation.pdf

References and Further Reading

Biesanz, J. C., Falk, C. F., & Savalei, V. (2010). Assessing mediational models: Testing and interval estimation for indirect effects. *Multivariate Behavioral Research, 45*, 661–701. http://dx.doi.org/10.1080/00273171.2010.498292
Chen, D., & Fritz, M. S. (2021). Comparing alternative corrections for bias in the bias-corrected bootstrap test of mediation. *Evaluation & The Health Professions, 44*(4), 416-427.
Judd, C.M. & Kenny, D.A. (1981). Process Analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5(5),* 602-619.
Falk, C. F., & Biesanz, J. C. (2015). Inference and interval estimation methods for indirect effects with latent variable models. *Structural Equation Modeling: A Multidisciplinary Journal, 22*(1), 24-38.
Fritz, M. S., & Mackinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18*, 233–239. http://dx.doi.org/10.1111/j.1467-9280.2007.01882.x
Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science,18*(3), 233-239.
Fritz, M. S., Taylor, A. B., & MacKinnon, D. P. (2012). Explanation of two anomalous results in statistical mediation analysis. *Multivariate Behavioral Research, 47*, 61-87.
Fritz, M. S., Kenny, D. A., & MacKinnon, D. P. (2016). The combined effects of measurement error and omitting confounders in the single-mediator model. Multivariate Behavioral Research, 51(5), 681-697.
Goodman, L. A. (1960). On the exact variance of products. *Journal of the American Statistical Association, 55*, 708-713.
Gonzalez, O., & MacKinnon, D. P. (2021). The measurement of the mediator and its influence on statistical mediation conclusions. *Psychological Methods, 26*(1), 1-17. https://doi.org/10.1037/met0000263
Kisbu-Sakarya, Y., MacKinnon, D. P., & Miočević, M. (2014). The distribution of the product explains normal theory mediation confidence interval estimation. *Multivariate Behavioral Research, 49*, 261-268. doi:10.1080/00273171.2014.903162
Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis.* New York: The Guilford Press.
Hoyle, R. H., & Kenny, D. A. (1999). Statistical power and tests of mediation. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research*. Newbury Park: Sage.
Kenny, D. A., & Judd, C. M. (2014). Power anomalies in testing mediation. *Psychological Science, 25*, 334–339.
Lachowicz, M. J., Preacher, K. J., & Kelley, K. (2018). A novel measure of effect size for mediation analysis. *Psychological Methods, 23*(2), 244-261.
MacKinnon, D.P. (2008). *Introduction to statistical mediation analysis.* Mahwah, NJ: Erlbaum.
MacKinnon, D. P., Cheong, J., Pirlott, A. G., & Smyth, H. L. (2023). Statistical mediation analysis in psychological research. In H. Cooper, M. N. Coutanche, L. M. McMullen, A. T. Panter, D. Rindskopf, and K. J. Sher, *APA handbook of research methods in psychology: Foundations, planning, measures, and psychometrics, Vol. 2.* (pp. 435-458). American Psychological Association.
MacKinnon, D.P. & Dwyer, J.H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review, 17(2*), 144-158.
MacKinnon, D.P., Fairchild, A.J., & Fritz, M.S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593-614.
MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*(4), 173–181. https://doi.org/10.1023/A:1026595011371
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.
MacKinnon DP, Lockwood CM, Williams J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*, 99–128
MacKinnon DP, Warsi G, Dwyer JH. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research, 30*:41–62
Miočević, M., O'Rourke, H. P., MacKinnon, D. P., & Brown, H. C. (2018). Statistical properties of four effect-size measures for mediation models. *Behavior research methods*, *50*, 285-301.

O'Rourke, H. P., & MacKinnon, D. P. (2015). When the test of mediation is more powerful than the test of the total effect. Behavior research methods, 47, 424-442. https://doi.org/10.3758/s13428-014-0481-z

O'Rourke, H. P., & MacKinnon, D. P. (2018). Reasons for testing mediation in the absence of an intervention effect: A research imperative in prevention and intervention research. *Journal of Studies on Alcohol and Drugs, 79*(2), 171–181. https://doi.org/10.15288/jsad.2018.79.171

Preacher, K.J., & Hayes, A.F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models.*Behavior Research Methods, Instruments, & Computers, 36*, 717-731.

Shrout, P.E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422-445.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological Methodology 1982* (pp. 290-312). Washington DC: American Sociological Association.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). *Mediation: R package for causal mediation analysis*. Retrieved from ftp://cran.r-project.org/pub/R/web/packages/mediation/vignettes/mediation.pdf

Tofighi, D., & MacKinnon, D. P. (2011). RMediation: An R package for mediation analysis confidence intervals. *Behavior Research Methods, 43*, 692-700.

Tofighi, D., & MacKinnon, D. P. (2016). Monte Carlo confidence intervals for complex functions of indirect effects. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(2), 194-205.

Valente, M. J., Gonzalez, O., Miočević, M., & MacKinnon, D. P. (2016). A note on testing mediated effects in structural equation models: Reconciling past and current research on the performance of the test of joint significance. *Educational and Psychological Measurement*, *76*(6), 889-911.

VanderWeele, T. J., Valeri, L., & Ogburn, E. L. (2012). The role of measurement error and misclassification in mediation analysis. *Epidemiology, 23*, 561–564.

Wang, C., & Xue, X. (2016). Power and sample size calculations for evaluating mediation effects in longitudinal studies. *Statistical Methods in Medical Research, 25*(2), 686-705.

Yzerbyt, V., Muller, D., Batailler, C., & Judd, C. M. (2018). New recommendations for testing indirect effects in mediational models: The need to report and test component paths. *Journal of Personality and Social Psychology, 115*(6), 929.