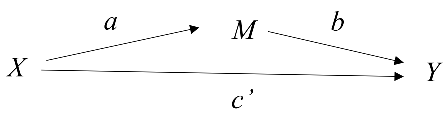
Mediation with Multilevel Regression

Mediation

A mediating relationship is a hypothesized causal chain of events that some predictor X predicts a another variable M which in turn predicts an outcome Y. Generally, mediator relationships involve a process by which some cause affects a consequence. An example might be that worker satisfaction impacts motivation which in turn impacts productivity. Or, high school students of higher SES may perform better on standardized tests and consequently get admitted to higher status colleges.



The mediational hypothesis can be tested by conducting a significance test of the mediational pathway, called an indirect effect. In multiple regression, the indirect effect is either the product of the regression coefficients a and b obtained from two regressions, one in which M is predicted by X and one in which Y is predicted by both X and M, or it is the difference between the coefficient for the simple regression with Y predicted by X, c, and the c' regression coefficient from the regression that includes the mediator, $c - c'$. In multiple regression, the product $a*b$ and the difference $c - c'$ are equivalent. The indirect regression coefficient represents the change in Y for every unit change in X as mediated by M. The standard error estimates for indirect effect coefficients are somewhat more complicated by the fact that there have been a number of proposed approaches. Percentile bootstrap and Monte Carlo appear to have the best power and Type I error balance (see Tofighi & MacKinnon, 2015).[1]

Multilevel Mediation

The tests of indirect effects in multilevel regression proceed in a similar fashion, although the two computational methods $a*b$ and the difference $c - c'$ will be similar but no longer exactly equivalent because of unequal group-size weighting (MacKinnon, 2008). With multilevel models, predictor variables can be from different levels. If all three variables are level-1 variables, then the mediation model is referred to as 1 → 1 → 1. If the X variable is from level 2 but the other two variables are level-1 variables, then it is a 2 → 1 → 1 mediation model (Krull & MacKinnon, 2001). 2 → 2 → 1, 1 → 2 → 1, and 1 → 1 → 2, for example, also are possible, although the latter two would have to involve one regression that was not truly a multilevel regression if the outcome from the regression is measured only at level 2. When the outcome (either M or Y) is measured at the individual level, the fact that β_{0j} serves as a dependent variable in level-2 of the multilevel regression equations suggests that it can be interpreted as a level-2 outcome. Where only level-2 measurement of the outcome variables exist, the indirect effect estimate would use the coefficient and standard errors from one or more single-level ordinary least squares regressions (Krull & MacKinnon, 2001). Random slopes can be included as theoretically and empirically appropriate. Caution is needed in interpreting these models, because the accuracy of the estimates, their statistical tests, and their interpretation depends on whether the appropriate level of measurement and corresponding centering approach is used in the model (Rockwood & Hayes, 2022; Zhang, Zyphur, & Preacher, 2009). Because level-1 measures may contain information about group context, effects may be due in part to causal processes occurring at level 2 unless corresponding level-2 measures of the constructs are also included in the model. For this reason, Zhang and colleagues suggest group-mean centering of level-1 variables when the level-1 process is of interest. Conversely, interpretation of variables measured at level 2 would not be appropriate if applied to causal processes that occur at level 1. For 1 → 1 → 1 models, Bauer and colleagues (Bauer, Preacher, & Gil, 2006) show that Monte Carlo and a normal approximation method using Kacker-Harville adjustment to the covariance of estimates provide generally accurate Type I error rates and coverage.

[1] See " Testing Mediation with Regression Analysis" handout from my multiple regression class page for an overview and references for mediation analysis.

Software Implementation

Software implementation for 1 → 1 → 1 models is available with the Monte Carlo method using Preacher and Selig's (2010) online calculation tool http://quantpsy.org/medmc/medmc111.htm. Nicholas Rockwood has a macro for SPSS for multilevel mediation, https://njrockwood.com/mlmed. Bauer and colleagues (Bauer, Preacher, & Gil, 2006) have a SAS macro for level-1 indirect effect standard errors and CIs. The `RMediation` package (Tofighi & MacKinnon, 2011) uses a distribution of products methods that allows for asymmetric sampling distribution of the indirect ($a*b$ product) coefficient, which performs well (Tofighi & MacKinnon, 2016). The R package `multilevelmediation` (Falk et al., 2024) supports random effects with 1-1-1 models but plans to add other cross-level mediation models, https://cran.r-project.org/web/packages/multilevelmediation/index.html. Mediation analysis with Kenny's actor-partner independence model (APIM) and be conducted with the help of an online tool, https://davidakenny.shinyapps.io/APIM_MM/. For more complex models that involve some level-2 variables (see Mathieu & Taylor, 2007, for some considerations), such as 2-1-1 or 2-2-1, the best option likely is to use a path analysis/structural equation modeling package that also estimates multilevel effects, such as Mplus (Muthén & Muthén, 1998-2017; see MacKinnon, 2010; Preacher, Zhang, & Zyphur, 2016: http://quantpsy.org/pubs/preacher_zhang_zyphur_2016_(code.appendix).pdf; Preacher, Zhang, & Zyphur, 2011; Zyphur and Zhang, 2010). Within this approach, Bayesian estimation approach in SEM (Yuan & MacKinnon, 2009) is possible and may produce better standard error estimates than maximum likelihood estimation if there are small number of groups (Zitzmann, Lüdtke, Robitzsch, 2015).

References

Bauer, D. J., Preacher, K. J. & Gil, K. M. (2006) Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods, 11*, 142-163.

Falk, C. F., Vogel, T. A., Hammami, S., & Miočević, M. (2024). Multilevel mediation analysis in R: A comparison of bootstrap and Bayesian approaches. *Behavior Research Methods, 56*(2), 750-764.

Kenny, D.A., Korchmaros, J. D., & Bolger, N. (2003). Lower level Mediation in multilevel models. *Psychological Methods, 8*, 115-128.

Kenny, D. A. (2015, February). An interactive tool for the estimation and testing the Actor-Partner Interdependence Model using multilevel modeling [Computer software]. Available from https://davidakenny.shinyapps.io/APIM_MM/.

Krull, J. L. & MacKinnon, D. P. (2001) Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research, 36*, 249-277.

MacKinnon, D. (2012). Chapter 10, *Introduction to statistical mediation analysis*. New York: Routledge.

Mathieu, J. E., & Taylor, S. R. (2007). A framework for testing meso-mediational relationships in Organizational Behavior. Journal of Organizational Behavior, 28, 141-172.

Muthén, L.K. and Muthén, B.O. (1998-2017). *Mplus User's Guide. Eighth Edition.* Los Angeles, CA: Muthén & Muthén

Preacher, K. J., & Selig, J. P. (2010, July). Monte Carlo method for assessing multilevel Mediation: An interactive tool for creating confidence intervals for indirect effects in 1-1-1 multilevel models [Computer software]. Available from http://quantpsy.org/.

Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures, 6*, 77-98.

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2016). Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychological methods, 21*, 189. Supplement with Mplus code: http://quantpsy.org/pubs/preacher_zhang_zyphur_2016_(code.appendix).pdf

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2011). Alternative methods for assessing mediation in multilevel data: The advantages of multilevel SEM. *Structural Equation Modeling, 18*, 161-182.

Rockwood, N.J. (2017). *Advancing the Formulation and Testing of Multilevel Mediation and Moderated Mediation Models*. Masters Thesis, Ohio State University.

Rockwood, N. J. & Hayes, A. F. (2017, May). *MLmed: An SPSS macro for multilevel mediation and conditional process analysis.* Poster presented at the annual meeting of the Association of Psychological Science (APS), Boston, MA. http://afhayes.com/public/aps2013.pdf. Macro: https://njrockwood.com/mlmed.

Rockwood, N.J., & Hayes, A.F. (2022). Multilevel mediation analysis. In A. A. O'Connell, D. B. McCoach, & B. A. Bell, (Eds.), *Multilevel modeling methods with introductory and advanced applications* (pp. 567-597). IAP.

Tofighi, D. and 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*, 194-205.

Tofighi, D., West, S. G., & MacKinnon, D. P. (2013). Multilevel mediation analysis: The effects of omitted variables in the 1–1–1 model. *British Journal of Mathematical and Statistical Psychology, 66*(2), 290-307.

Zitzmann, S., Lüdtke, O., & Robitzsch, A. (2015). A Bayesian approach to more stable estimates of group-level effects in contextual studies. Multivariate behavioral research, 50(6), 688-705.

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods, 12*, 695-719.