Current Directions in Mediation Analysis

David P. MacKinnon[1] and Amanda J. Fairchild[2]

[1]*Arizona State University and* [2]*University of South Carolina*

ABSTRACT—*Mediating variables continue to play an important role in psychological theory and research. A mediating variable transmits the effect of an antecedent variable on to a dependent variable, thereby providing more detailed understanding of relations among variables. Methods to assess mediation have been an active area of research for the last two decades. This paper describes the current state of methods to investigate mediating variables.*

KEYWORDS—*mediation, indirect effect, statistical methods*

Mediating variables have been ubiquitous throughout the history of psychology because they are used to explain how or why two variables are related. One of the oldest models in psychology, the Stimulus → Organism → Response model, specifies mediating processes in an organism that transmit a stimulus to a response. A modern application of mediating variables is in treatment and prevention research, where interventions are designed to change mediating variables, such as norms, which are hypothesized to be causally related to a dependent variable, such as health behavior. The promise of mediation analysis is that it can identify fundamental processes underlying human behavior that are relevant across behaviors and contexts. Once a true mediating process is identified, then more efficient and powerful interventions can be developed because these interventions can focus on variables in the mediating process. A variety of mediation-analysis methods, including statistical and experimental methods, have been used throughout the history of psychology. New developments in mediation analysis extract more accurate information about whether a variable truly mediates the relation between two other variables.

Consider some examples of mediating variables in psychology: (a) A tobacco prevention program, the antecedent variable, reduces cigarette smoking by changing the social norms for tobacco use, the mediating variable; (b) exposure to negative life events affects blood pressure through the mediation of cognitive attributions to stress; or (c) instructing persons to make images of words in a memory task increases the number of images made, the mediating variable, which improves recall of the words. In each case, an antecedent variable affects a mediator variable and the mediator variable affects a dependent variable, thus forming a chain of relations among the three variables (Baron & Kenny, 1986; Judd & Kenny, 1981). The chain of relations among the variables is called an indirect or mediated effect of the antecedent variable on the dependent variable. An effect that is not mediated this way is called a direct effect. Although the consideration of a mediating variable may appear straightforward, even this simplest of mediating-variable models can be complicated. The methodological and statistical challenges of investigating mediation have made the assessment of mediation an active research topic.

THE SINGLE-MEDIATOR MODEL

The single-mediator model is shown in Figure 1 and further illustrated with the following three equations (MacKinnon & Dwyer, 1993):

$$(1) \mathrm{Y} = i_1 + c\mathrm{X} + e_1$$

$$(2) \mathrm{Y} = i_2 + c'\mathrm{X} + b\mathrm{M} + e_2$$

$$(3) \mathrm{M} = i_3 + a\mathrm{X} + e_3$$

where Y is the dependent variable, X is the antecedent variable, and M is the mediating variable. The coefficient, c, represents how strongly X predicts Y; c' represents the strength of prediction of Y from X, with the strength of the M-to-Y relation removed; b is the coefficient for the strength of the relation between M and Y with the strength of the X-to-Y relation removed; and a is the coefficient representing the strength of the relationship between X and M. The intercepts in each equation, representing the average score of each variable, are i_1 and i_2 and i_3, respectively; and e_1, e_2, and e_3 represent the error, or the part of the relation that cannot be predicted.

Address correspondence to David P. MacKinnon, Department of Psychology, Arizona State University, Tempe, AZ 85287-1104; e-mail: davidpm@asu.edu.
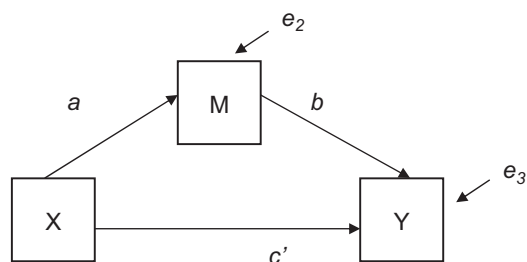
Fig. 1. The single-mediator model. The relation between the antecedent variable X on the dependent variable Y is mediated through variable M; *a* represents the strength of the relation between X and M, *b* represents the strength of the relation between M and Y, *c′* represents the strength of the relation between X and Y, e_2 represents the unexplained part of M, and e_3 represents the unexplained part of Y.

Several different effects are represented in the model: (a) a direct effect relating X to Y with the strength of mediator relation removed, quantified by *c′*; (b) a mediated, or indirect, effect of X to Y transmitted through the mediating variable, quantified by *ab* (or *c − c′*); and (c) a total effect of X on Y that is computed by the addition of these two parts. The numerical values of the mediated effect may be computed in one of two ways: as either the difference in coefficients $\hat{c} - \hat{c}'$ (a hat above a coefficient, e.g., $\hat{c}$, indicates that it is an estimate in a sample of data) or as the product of coefficients $\hat{a}\hat{b}$ (MacKinnon & Dwyer, 1993).

SIGNIFICANCE TESTING OF THE MEDIATED EFFECT

There have been several methods proposed for assessing whether a mediated effect is large enough to be considered important, which is known in the statistical literature as a test of statistical significance. Each of these methods uses at least two of Equations 1, 2, and 3.

Causal Steps

The causal-steps approach to testing mediation was outlined in Baron and Kenny (1986) and Judd and Kenny (1981) and has been the most commonly used approach to test mediation. Generally, the method requires separate significance tests of the strength of the overall relation between X and Y (*c*), the strength of the relation between X and M (*a*), the strength of the relation between M and Y adjusted for X (*b*), and visual inspection of whether $\hat{c}$ is greater than $\hat{c}'$. Despite its wide use in substantive research, there are several limitations to the causal-steps approach. First, recent statistical-simulation studies show that the ability to detect mediated effects using the causal-step method can be very low (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). A recent study showed that the Baron and Kenny causal-steps approach required approximately 21,000 subjects for adequate ability to detect an effect when the effect sizes of the *a* and *b* paths were of small strength and all of the relation of X to Y was mediated (Fritz & MacKinnon, 2007). Second, the method

does not explicitly provide a numerical value of the strength of the mediated effect. Finally, the test requires that there be a significant overall relation between X and Y for mediation to exist.

The requirement of a significant overall relation between X and Y is the central difference between the causal-steps approach and other methods for testing mediation. Some researchers have treated this test of the overall relation between X and Y as a perfect test of the relation, failing to recognize that it is a fallible statistical test that is subject to error, and arguing that if there is not a significant overall effect then mediation should not be examined. The requirement that X is significantly related to Y is an important test in any research study, but mediation can exist even in the absence of such a significant relation. The statistical test of the effect of X on Y can have less power than the test of the links in the mediation model. Several scenarios illustrate where significant mediation exists but the overall effect of X on Y is not significant. Consider a case of mediation in which there are subgroups of individuals for whom the mediated effect is of opposite sign (i.e., positive versus negative), such that a test of the X-to-Y relation for the pooled data would be zero even though mediation exists in the data. Or consider a case in which the sign of the mediated effect, *ab*, differs from the sign of the direct effect (*c′*), causing the overall relation of X to Y (*c*) to be zero (such cases are known as inconsistent-mediation models).

Estimates of the Mediated Effect

There are formulas for the variability of difference in coefficients $\hat{c} - \hat{c}'$ and for the product of coefficients $\hat{a}\hat{b}$ (MacKinnon et al., 2002). The mediated effect can be tested for statistical significance by dividing the estimate by its standard error (square root of the variance) and comparing the ratio to the standard normal distribution, although there are better tests of significance, as we will describe.

CONFIDENCE LIMITS FOR THE MEDIATED EFFECT

In addition to testing the mediated effect, $\hat{a}\hat{b}$ (or $\hat{c} - \hat{c}'$), for significance, limits for the true value of *ab* can be constructed. The confidence limits for the mediated effect provide information on the reliability or accuracy of the estimate of the mediated effect. Recent research has shown that confidence limits and significance testing for the mediated effect based on the normal distribution are often inaccurate and are unlikely to find a real mediated effect in a sample of data (MacKinnon et al., 2002). Asymmetric confidence limits based on the distribution of the product, *ab*, and methods based on repeatedly sampling the original data are more accurate. These tests capture the non-normal shape of the mediated-effect sampling distribution (which occurs because the strength of the mediated effect is the product of two coefficients and does not always have a normal distribution), thus improving power. A new downloadable pro-

gram, PRODCLIN (MacKinnon, Fritz, Williams, & Lockwood, 2007), constructs confidence limits for the mediated effect based on the distribution of the product. Resampling methods are another option to handle the non-normality in the distribution of the mediated effect (Shrout & Bolger, 2002).

ASSUMPTIONS OF THE SINGLE-MEDIATOR MODEL

Several assumptions underlie these tests of mediation, including assumptions necessary for the statistical methods used to estimate the strength of relations in the regression models. Assumptions of the correct specification of causal ordering and causal direction are also especially important but often difficult to defend. New approaches have been developed to improve causal inference from the mediation model based on consideration of counterfactual experimental conditions in which participants could have theoretically served in addition to conditions in which the participants actually did serve. These causal-inference models expose hidden assumptions of mediation models and, in some cases, suggest methods to address violations of model assumptions (Frangakis & Rubin, 2002). A strength of causal-inference approaches is that they clarify additional criteria to establish causal mediation (MacKinnon, 2008).

LONGITUDINAL MEDIATION MODELS

Longitudinal mediation models permit the examination of several mediation questions that cannot be asked using the cross-sectional mediation model, such as whether a mediated effect is stable over time. Longitudinal models also shed light on temporal-precedence or causal-ordering assumptions by quantifying mediation relations among variables over time. There are three major types of longitudinal mediation models: (a) the autoregressive model, (b) latent-growth models, and (c) latent-difference-score models. In the basic autoregressive model, dependency between adjacent longitudinal relations is specified and relations consistent with longitudinal mediation are estimated. Cole and Maxwell (2003) provide a detailed description of the many longitudinal mediation relations for an autoregressive model with multiple measurement occasions.

The latent-growth mediation model examines whether growth in an independent variable, X, affects the growth of a mediator, M, which affects the growth of an outcome, Y. The latent-difference-score model specifies differences between waves of observations, such as change in X affecting later change in M, which affects later change in Y. Both the latent-growth and latent-difference-score models use latent variables to represent growth or change over time. In all of the longitudinal models, measurement invariance is critical because changes in measurement over time confound the interpretation of change over time.

MODELS THAT INCLUDE MODERATION AND MEDIATION

Several different models to estimate the strength of effects that may occur when mediation and moderation analyses are combined have been proposed (e.g., Edwards & Lambert, 2007; MacKinnon, 2008). First, the strength and/or form of a mediated effect may depend on a moderator variable. This effect has been termed "moderated mediation" (e.g., Baron & Kenny, 1986). Examining whether a mediated effect is constant across levels of a moderator variable provides a means to examine the generalizability of the mediated relation. Investigations such as these may be particularly useful in assessing whether a program achieves its effects in similar ways across subgroups of interest, such as naturally occurring groups like race or gender or experimentally manipulated groups like treatment or control.

Second, the effect of an interaction on the dependent variable may be transmitted through a mediator, such that the mediator variable is intermediate in the causal sequence from an interaction effect to a dependent variable. This effect has been termed "mediated moderation" (e.g., Baron & Kenny, 1986). The purpose of this analysis is to determine the mediating variable(s) that explain the interaction effect. For example, the effect of a drug prevention program may be greater for high-risk subjects because the social norm for this group may change more than it does for low-risk subjects.

EXPERIMENTAL-DESIGN APPROACHES TO ASSESSING MEDIATION

Statistical testing of mediation relations may also be combined with experimental design. Here, participants are randomized to levels of one or more factors in order to demonstrate a pattern of results consistent with one mediation theory and inconsistent with another theory (MacKinnon, 2008). Statistical mediation analysis is conducted using one or several antecedent variables reflecting main and interaction effects. One type of mediation study is called a blockage design, in which the manipulation is designed to remove the effects of a mediating variable so that the mediation relation is observed in one experimental condition but not in the condition where it is blocked. Another design, called an enhancement design, would randomly assign participants to a condition that would enhance effects of the mediator so that the mediation relation is larger in the enhanced condition than another condition. A recent paper argues for a more elaborate mediation experimental design whereby separate randomized experiments are conducted to investigate both the X-to-M relation and the M-to-Y relation (Spencer, Zanna, & Fong, 2005). One complicating aspect of this type of study is that the manipulation of the mediator in the M-to-Y experiment requires a mediation analysis as well because M may not be perfectly manipulated. The combination of experimental design with statistical mediation analysis is a very promising

methodology because it combines the interpretability of randomized manipulations with estimation of mediation.

OTHER DIRECTIONS IN MEDIATION ANALYSIS

There have been other extensions to mediation analysis in addition to the ones already described. Methods for testing mediation models with dependent variables that have two categories (e.g., depressed or not, smoked or not, or littered or not) have been outlined and compared (e.g., MacKinnon, Lockwood, Brown, Wang, & Hoffman, 2007), and models that accommodate multiple mediators have been developed and evaluated (Taylor, MacKinnon, & Tein, 2008). Mediation models for data collected from persons in clinics or schools (multilevel data) have also been developed, including cases in which participants have been repeatedly measured (e.g., Bauer, Preacher, & Gil, 2006).

FUTURE DIRECTIONS

Mediation analysis continues to be an important area of substantive and methodological research. It is likely that longitudinal mediation models will be an active area of development and application because of the importance of longitudinal data for investigation of temporal precedence in mediation relations. Alternative longitudinal models will be compared analytically, in statistical simulations, and in application to real data. Substantive theory is needed to address the process by which variables change over time and how the change process between X and M may be related to the change process between M and Y, adjusted for X, including situations in which the types of change in these two mediation relations differ. Models for change over time may include cascades, threshold, and cumulative models. Models for assessing change in continuous time have the promise of specifying mediation relations applicable to any data-collection time frame.

Another future direction regards methods to draw causal conclusions from mediation analysis. Future research will compare different approaches for assessing causal inference especially when relations among variables are not based on randomization. Methods to test violations of assumptions and methods to address how these violations affect conclusions from a mediation analysis will be developed. Similarly, the best experimental designs and programs of research to more thoroughly investigate mediation are needed. Models that incorporate moderation and mediation relations will be refined and extended for longitudinal data, different distributions of variables, and mixtures of individuals. The promise of these models is that mediation processes that depend on other variables may be more clearly identified. All of the existing and forthcoming developments in mediation analysis will help obtain more accurate answers to the ubiquitous question in psychology of how and why two variables are related to each other.

Recommended Reading

Holland, P.W. (1988). Causal inference, path analysis, and recursive structural equation models. *Sociological Methodology*, *18*, 449–484. A complex but thorough examination of causal inference for the mediation model.

Judd, C.M., & Kenny, D.A. (1981). (See References). A classic paper on mediation that includes original ideas on many topics such as longitudinal mediation and moderated mediation that have been refined since its publication.

MacKinnon, D.P. (2008). (See References). A general introduction to the current state of mediation analysis.

MacKinnon, D.P., Fairchild, A.J., & Fritz, M.S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614. A review of mediation analysis in psychology.

Sobel, M.E. (1986). Some new results on indirect effects and their standard errors in covariance structure models. *Sociological Methodology*, *13*, 290–312. Describes a general approach to finding the standard errors of indirect effects in simple and complex models that is now used in major covariance-structure-analysis software.

REFERENCES

Baron, R.M., & Kenny, D.A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.

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.

Cole, D.A., & Maxwell, S.E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*, 558–577.

Edwards, J.R., & Lambert, L.S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods*, *12*, 1–22.

Frangakis, C.E., & Rubin, D.B. (2002). Principal stratification in causal inference. *Biometrics*, *58*, 21–29.

Fritz, M.S., & MacKinnon, D.P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, *18*, 233–239.

Judd, C.M., & Kenny, D.A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, 602–619.

MacKinnon, D.P. (2008). *Introduction to statistical mediation analysis*. Mahwah, NJ: Erlbaum.

MacKinnon, D.P., & Dwyer, J.H. (1993). Estimation of mediated effects in prevention studies. *Evaluation Review*, *17*, 144–158.

MacKinnon, D.P., Fritz, M.S., Williams, J., & Lockwood, C.M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behavior Research Methods*, *39*, 384-389. (The program can be downloaded from http://www.public.asu.edu/ ∼ davidpm/ripl/Prodclin/)

MacKinnon, D.P., Lockwood, C.M., Brown, H., Wang, W., & Hoffman, J. (2007). The intermediate endpoint effect in logistic and probit regression. *Clinical Trials*, *4*, 499–513.

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.

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

Spencer, S.J., Zanna, M.P., & Fong, G.T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, *89*, 845–851.

Taylor, A.B., MacKinnon, D.P., & Tein, J.Y. (2008). Tests of the three-path mediated effect. *Organizational Research Methods*, *11*, 241–269.