Prev Sci (2009) 10:87–99
DOI 10.1007/s11121-008-0109-6

A General Model for Testing Mediation and Moderation Effects

Amanda J. Fairchild · David P. MacKinnon

Published online: 12 November 2008

Abstract This paper describes methods for testing mediation and moderation effects in a dataset, both together and separately. Investigations of this kind are especially valuable in prevention research to obtain information on the process by which a program achieves its effects and whether the program is effective for subgroups of individuals. A general model that simultaneously estimates mediation and moderation effects is presented, and the utility of combining the effects into a single model is described. Possible effects of interest in the model are explained, as are statistical methods to assess these effects. The methods are further illustrated in a hypothetical prevention program example.

Keywords Mediation · Indirect effect · Moderation · Mediated moderation · Moderated mediation

Relations between variables are often more complex than simple bivariate relations between a predictor and a criterion. Rather these relations may be modified by, or informed by, the addition of a third variable in the research design. Examples of third variables include suppressors, confounders, covariates, mediators, and moderators (MacKinnon et al. 2000). Many of these third variable effects have been investigated in the research literature, and more recent research has examined the influences of more than one third variable effect in an analysis. The importance of investigating mediation and moderation effects together has been recognized for some time in prevention science, but statistical methods to conduct these analyses are only now being developed. Investigations of this kind are especially valuable in prevention research where data may present several mediation and moderation relations.

Previous research has described the differences between mediation and moderation and has provided methods to analyze them separately (e.g., Dearing and Hamilton 2006; Frazier et al. 2004; Gogineni et al. 1995; Rose et al. 2004). More recent research has presented models to simultaneously estimate mediation and moderation to investigate how the effects work together (e.g., Edwards and Lambert 2007; MacKinnon 2008; Muller et al. 2005; Preacher et al. 2007). A review of the substantive literature illustrates that few applied research examples have used these models, however. Although analyzing mediation and moderation separately for the same data may be useful, as described later in this paper, simultaneous examination of the effects is often relevant and allows for the investigation of more varied, complex research hypotheses.

What Type of Research Questions Can Be Addressed with the Simultaneous Analysis of Mediation and Moderation Effects?

"Is the Process By Which a Program Has an Effect the Same Across Different Types of Participants?"

In prevention and intervention research, the mediation model has been used to understand the mechanism(s) by which program effects occur. To determine the generalizability

A. J. Fairchild (✉)
Department of Psychology, University of South Carolina,
Barnwell College,
1512 Pendleton St.,
Columbia, SC 29208, USA
e-mail: afairchi@mailbox.sc.edu

D. P. MacKinnon
Research in Prevention Lab, Department of Psychology,
Arizona State University,
P.O. Box 871104, Tempe, AZ 85287-1104, USA

of these mechanisms or to explain an unexpectedly small mediated effect it may be of interest to investigate whether the mediation relation, or the indirect effect, holds across different subgroups (e.g., men vs. women or low-risk vs. high-risk). To investigate these hypotheses, a researcher asks whether the indirect effect is moderated, or whether the mediated effect depends on levels of another variable. For example, suppose that a business implements a worksite-wellness program (the independent variable, X) to reduce obesity-related health risks in its employees. Program developers hypothesize that by increasing employee knowledge about the benefits of eating fruits and vegetables (the mediator variable, M), employee consumption of fruits and vegetables will increase (the dependent variable, Y), thus reducing health risk. An estimate of the indirect effect of the program on employee fruit and vegetable consumption through employee knowledge of the benefits of eating fruits and vegetables is unexpectedly low. Through talks with employees, it becomes apparent that participants were more or less motivated to gain and use knowledge from the program to improve their diet based on whether they had a family history of obesity-related illness such as diabetes or cardiovascular disease. Program developers hypothesize that participants' family history of obesity-related illness may moderate the mediation relation in the data, affecting the influence of the program on employee knowledge of fruits and vegetables and its subsequent impact on fruit and vegetable consumption (See Fig. 1).

"Can a Mediation Relation Explain an Interaction Effect in My Data?"

Suppose a similar worksite-wellness program was implemented in a larger sister company and program effects had been dependent on whether the participant was a full or part-time employee at the company. To investigate the underlying reasons for this unexpected interaction, or moderation relation, program analysts could investigate a mediation hypothesis where the interaction effect predicts a mediator variable which predicts the outcome, defined here as *the mediation of a moderator effect*. For example, perhaps in addition to increasing employee knowledge of fruit and vegetable benefits with the wellness curriculum, the program (X) also introduced a work culture, or a social norm (M), of healthy eating which contributed to employee fruit and vegetable consumption (Y; See Fig. 2). Program developers hypothesize that the more hours an employee worked in a week determined how much they were subjected to the social norm which ultimately influenced their fruit and vegetable consumption.

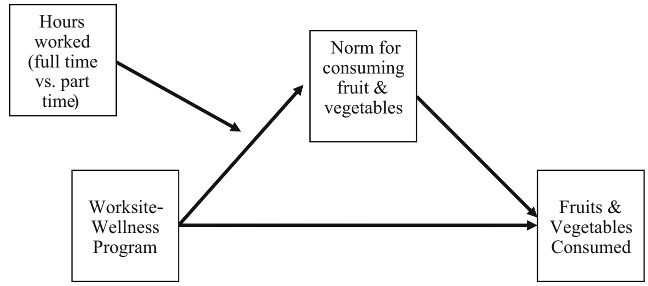

Fig. 2 Conceptual diagram for the mediation of a moderator effect example

Current Research

The purpose of this article is to provide a straightforward, methodological resource on models to simultaneously test mediation and moderation effects for the substantive researcher. To that end, we organize methods for simultaneously testing mediation and moderation into a single framework that allows for point estimation and construction of confidence intervals. Interpretation and effect computation are provided, and the model is applied to a substantive dataset to illustrate the methods. To ensure common ground for this discussion, basic mediation and moderation effects from which the model is formed are first reviewed.

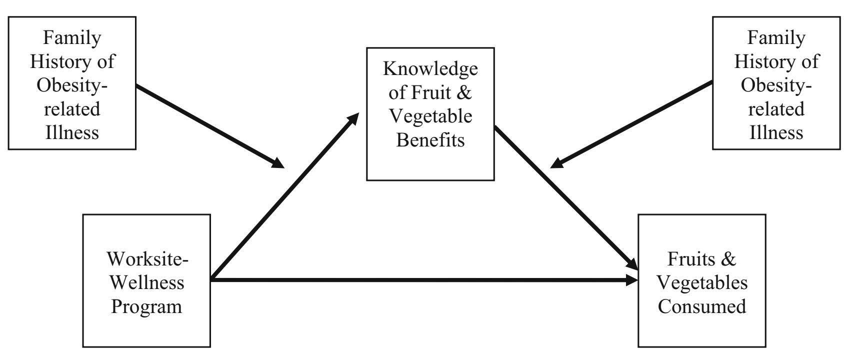

Fig. 1 Conceptual diagram for the moderation of an indirect effect example

Review of the Mediation Model

The mediation model offers an explanation for how, or why, two variables are related, where an intervening or mediating variable, M, is hypothesized to be intermediate in the relation between an independent variable, X, and an outcome, Y (See Fig. 3). Early presentations of mediation in prevention research (e.g., Baron and Kenny 1986; Judd and Kenny 1981a; 1981b) illustrated causal step methods to test for mediation, but more recent research has supported tests for statistical mediation based on coefficients from two or more of the following regression equations (MacKinnon and Dwyer 1993):

$$Y = i_1 + cX + e_1 \tag{1}$$

$$Y = i_2 + c'X + bM + e_2 \tag{2}$$

$$M = i_3 + aX + e_3 \tag{3}$$

Where c is the overall effect of the independent variable on Y; c' is the effect of the independent variable on Y controlling for M; b is the effect of the mediating variable on Y; a is the effect of the independent variable on the mediator; i_1, i_2 , and i_3 are the intercepts for each equation; and e_1, e_2, and e_3 are the corresponding residuals in each equation (see Fig. 3).

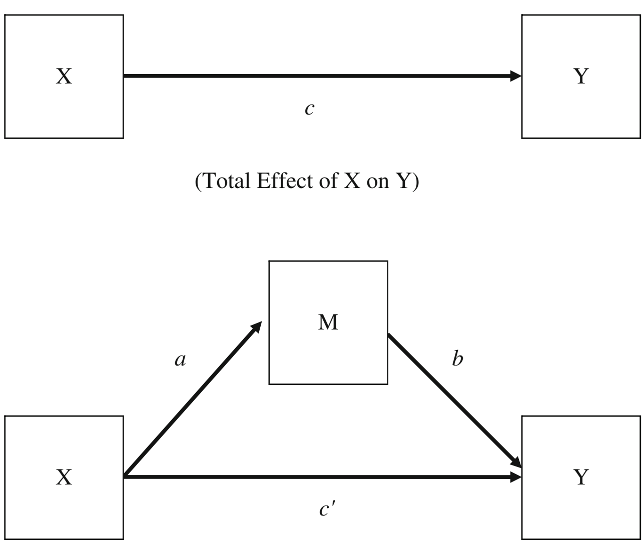

Fig. 3 Path diagram for the single-mediator model. *Note.* X= the independent variable, Y= the dependent variable, and M= the mediating variable. The mediation model decomposes the total effect of X on Y (c), into two parts: the indirect effect of X on Y, quantified by ab (the product of a and b), and the direct effect of X on Y with the effect of the mediator removed, quantified by c'. $c = ab + c'$

Although there are alternative ways to estimate mediation, the product of coefficients is most easily applied to complex models and is used in this paper. The product of coefficients test computes the mediated effect as the product of the $\hat{a}$ *and* $\hat{b}$ coefficients from Eq. 2 and 3. Sobel (1982, 1986) derived the variance of $\hat{a}\hat{b}$ product based on the multivariate delta method. This formula has been widely used to estimate the normal theory standard error of $\hat{a}\hat{b}$:

$$s_{\hat{a}\hat{b}} = \sqrt{s_{\hat{a}}^2 \hat{b}^2 + s_{\hat{b}}^2 \hat{a}^2} \tag{4}$$

Where $s_{\hat{a}}^2$ is the variance of the $\hat{a}$ coefficient and $S_{\hat{b}}^2$ is the variance of the $\hat{b}$ coefficient.

MacKinnon et al. (1998) and MacKinnon and Lockwood (2001) showed that tests for the mediated effect based on normal theory can yield inaccurate confidence limits and significance tests, however, as the product of two normally distributed variables is not itself normally distributed. Alternative tests based on the asymmetric distribution of the product of two normally distributed variables are available and have been shown to outperform traditional methods (MacKinnon et al. 2002; MacKinnon et al. 2004). A new program called "PRODCLIN" (MacKinnon et al. 2007) has automated computation of the distribution of the product test for mediation so that it is widely accessible. The researcher need only specify values of $\hat{a}$, $\hat{b}$, the standard error of $\hat{a}$, the standard error of $\hat{b}$, and the statistical significance level desired.

Assumptions of the mediation model include the usual OLS estimation assumptions (e.g., correct specification of the model's functional form, no omitted variables, no measurement error; Cohen et al. 2003). Mediation analysis also assumes correct causal ordering of the variables, no reverse causality effects, and no XM interaction.

Review of the Moderation Model

The moderation model tests whether the prediction of a dependent variable, Y, from an independent variable, X, differs across levels of a third variable, Z (See Fig. 4). Moderator variables affect the strength and/or direction of the relation between a predictor and an outcome: enhancing, reducing, or changing the influence of the predictor. Moderation effects are typically discussed as an interaction between factors or variables, where the effects of one variable depend on levels of the other variable in analysis. Detailed descriptions of moderator effects and a framework for their estimation and interpretation were presented in Aiken and West (1991).

Moderation effects are tested with multiple regression analysis, where all predictor variables and their interaction

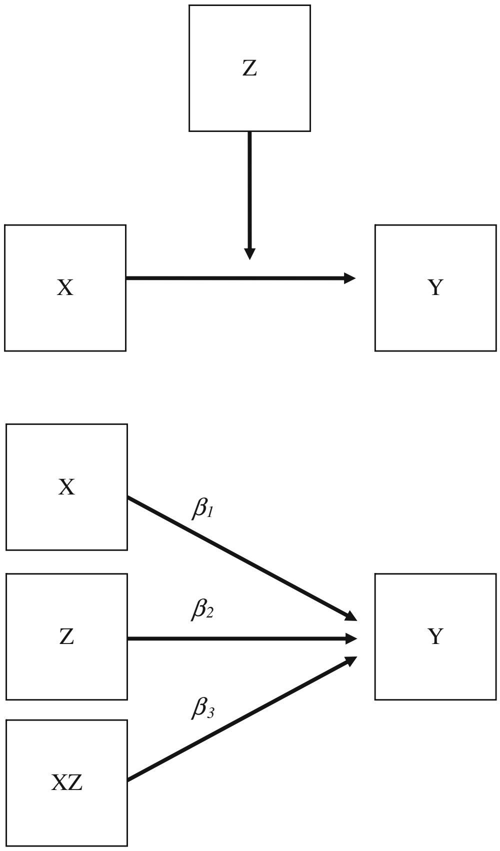

Fig. 4 Alternate path diagram representations of the moderation model. *Note.* X= the independent variable, Y= the dependent variable, Z= the moderator variable, XZ= the product of X and the moderator variable, β_1 = the effect of X on Y, β_2 = the effect of Z on Y, and β_3 = the effect of XZ on Y

term are centered prior to model estimation to improve interpretation of regression coefficients. A single regression equation forms the basic moderation model:

$$Y = i_5 + \beta_1 X + \beta_2 Z + \beta_3 XZ + e_5 \tag{5}$$

Where β_1 is the coefficient relating the independent variable, X, to the outcome, Y, when Z = 0, β_2 is the coefficient relating the moderator variable, Z, to the outcome when X = 0, i_5 the intercept in the equation, and e_5 is the residual in the equation.

The regression coefficient for the interaction term, β_3, provides an estimate of the moderation effect. If β_3 is statistically different from zero, there is significant moderation of the X-Y relation in the data. Plotting interaction effects aids in the interpretation of moderation to show how the slope of Y on X is dependent on the value of the moderator variable. Regression slopes that correspond to the prediction of Y from X at a single value of Z are termed simple slopes.

Assumptions of the moderation model include OLS regression assumptions, as described earlier, and homogeneity of error variance. The latter assumption requires that the residual variance in the outcome that remains after predicting Y from X is equivalent across values of the moderator variable.

Combining Mediation and Moderation Analyses

Analyzing the Models Separately

Much of the work combining mediation and moderation analyses has been presented in the context of prevention program design and development, where examining mediation and moderation effects together aims to improve program implementation by combining theory-driven ideas and empirical evidence. For example, Donaldson (2001) indicates that multivariate relations between variables in a treatment program tend to be one of three types: (a) direct effects, (b) mediated effects, and (c) moderated effects. By combining the examination of these effects in a single analysis, the researcher may not only identify mediating processes through which the program achieves its effects but may also identify effective program components and/or particular characteristics of the participants or the environment that moderate the effectiveness of the program. If the theoretical underpinnings of a treatment or prevention program serve as a starting point for its curriculum, separate analyses of mediation and moderation may be used to iteratively refine program theory. These analyses may be used to collect empirical feedback and to conduct pilot work of the program before large-scale implementation of the curriculum (See Fig. 5). Specifically, by examining mediation one is able to investigate how effective a program curriculum was in changing target behaviors, and whether the program aimed to alter appropriate mediators of desired outcomes. Analyzing moderation effects in this context allows the researcher to identify variables that may improve or reduce the program's ability to alter mediating variables, as well as to examine the external validity, or

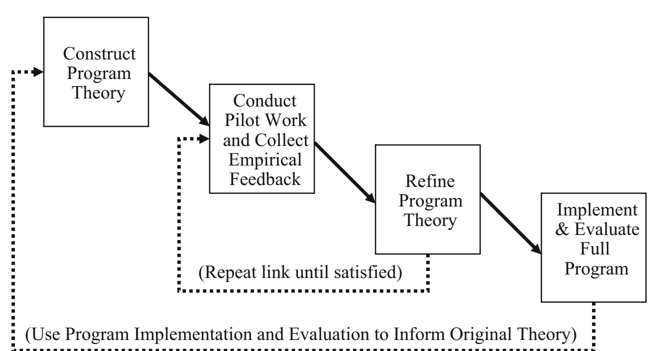

Fig. 5 Refining program theory: an empirical-theoretical exchange. *Note.* Figure is based on Donaldson (2001) diagram that occurs on p. 472 of that text

generalizability, of the model across different groups or settings (Hoyle and Robinson 2003). Hypothesized moderator variables may be more or less amenable to program tailoring, however. Although program subgroups may be formed on moderators such as age or gender with little difficulty, forming program subgroups based on other moderator variables such as ethnicity or family risk may be impractical and/or unethical. Nonetheless, the identification of subgroups for which a program is most effective is useful, and the examination of moderation and mediation effects in this context increases the scientific understanding of behaviors and improves program efficacy. West and Aiken (1997) have argued that these analyses are especially useful after the successful implementation and evaluation of a treatment program. This allows for the continual development and improvement of a program, but after an effective first evaluation.

Analyzing the Models Simultaneously

By simultaneously investigating mediation and moderation, the effects may not only be disentangled and analyzed separately but can also be evaluated together. There have been two primary effects analyzed in the literature: (a) the mediation of a moderator effect, and (b) the moderation of an indirect effect. The mediation of a moderator effect involves exploring mediating mechanisms to explain an overall interaction of XZ in predicting Y, whereas the moderation of an indirect effect involves investigating whether a mediated relation holds across levels of a fourth, moderating variable. These effects have previously been referred to as mediated-moderation and moderated-mediation in the literature, respectively. These alternative descriptions may enhance the distinction between the two.

Previous models to simultaneously test mediation and moderation effects have been presented with varying notation (e.g., Edwards and Lambert 2007; James and Brett 1984; Muller et al. 2005; Preacher et al. 2007) or without testable equations (e.g., Baron and Kenny 1986; Wegener and Fabrigar 2000), making it difficult to understand similarities and differences among the methods. Moreover the criteria for testing the effects have varied across sources, making it hard to extrapolate recommendations for use. It is possible to create a general model to test these effects, however, that subsumes several previous frameworks by including all possible interactions between variables in the mediation and moderation models (MacKinnon 2008). Such a model unifies the methods into a single presentation where different models are represented as special cases of the larger framework. Three regression equations form the model:

$$Y = i_6 + c_1 X + c_2 Z + c_3 XZ + e_6 \tag{6}$$

$$M = i_7 + a_1 X + a_2 Z + a_3 XZ + e_7 \tag{7}$$

$$\begin{aligned} Y = i_8 + c'_1 X + c'_2 Z + c'_3 XZ + b_1 M + b_2 MZ \\ + hXM + jXMZ + e_8 \end{aligned} \tag{8}$$

where all predictors in the model are centered at zero to improve interpretation of the lower order coefficients. In Eq. 6, c_1 is the effect of the independent variable on the outcome when $Z = 0$ (also the average effect of X on Y because the mean of $Z = 0$), c_2 is the effect of the moderator variable on the outcome when $X = 0$ (also the average effect of Z on Y because the mean of $X = 0$), c_3 is the effect of the interaction between the independent variable and the moderator on the outcome, and i_6 and e_6 are the intercept and the residual in the equation, respectively. In Eq. 7, a_1 is the effect of the independent variable on the mediator when $Z = 0$ (also the average effect of X on M because the mean of $Z = 0$), a_2 is the effect of the moderator variable on the mediator (also the average effect of Z on M because the mean of $X = 0$), a_3 is the effect of the interaction between the independent and moderator variables on the mediator, and i_7 and e_7 are the intercept and the residual in the equation, respectively. In Eq. 8, c'_1 is the effect of the independent variable on the outcome when $M = 0$ and $Z = 0$ (the average effect of X on Y), c'_2 is the effect of the moderator on the outcome when $X = 0$ and $M = 0$ (the average effect of Z on Y), c'_3 is the effect of the interaction between the independent and moderator variables on the outcome when $M = 0$ (the average effect of XZ on Y), b_1 is the effect of the mediator on the outcome when $X = 0$ and $Z = 0$ (the average effect of M on Y), b_2 is the effect of the interaction between the moderator and mediator variables on the outcome when $X = 0$ (the average effect of MZ on Y), h is the effect of the interaction between the independent and mediator variables on the outcome when $Z = 0$ (the average effect of XM on Y), and j is the effect of the three-way interaction of the mediating, moderating, and independent variables on the outcome. The intercept and residual in Eq. 8 are coded i_8 and e_8, respectively. A path diagram for the model is presented in Fig. 6.

Assumptions of the general model include assumptions of the mediation and moderation models as described earlier. Issues of causal inference in non-additive models may also require additional stipulations for estimation. Note that the presence of any significant two-way interactions in the model implies that the main effects of X and M do not provide a complete interpretation of effects. The presence of a significant three-way interaction in the model also implies that lower order two-way interactions do not provide a complete interpretation of effects. If there are significant interactions, point estimates can be probed with

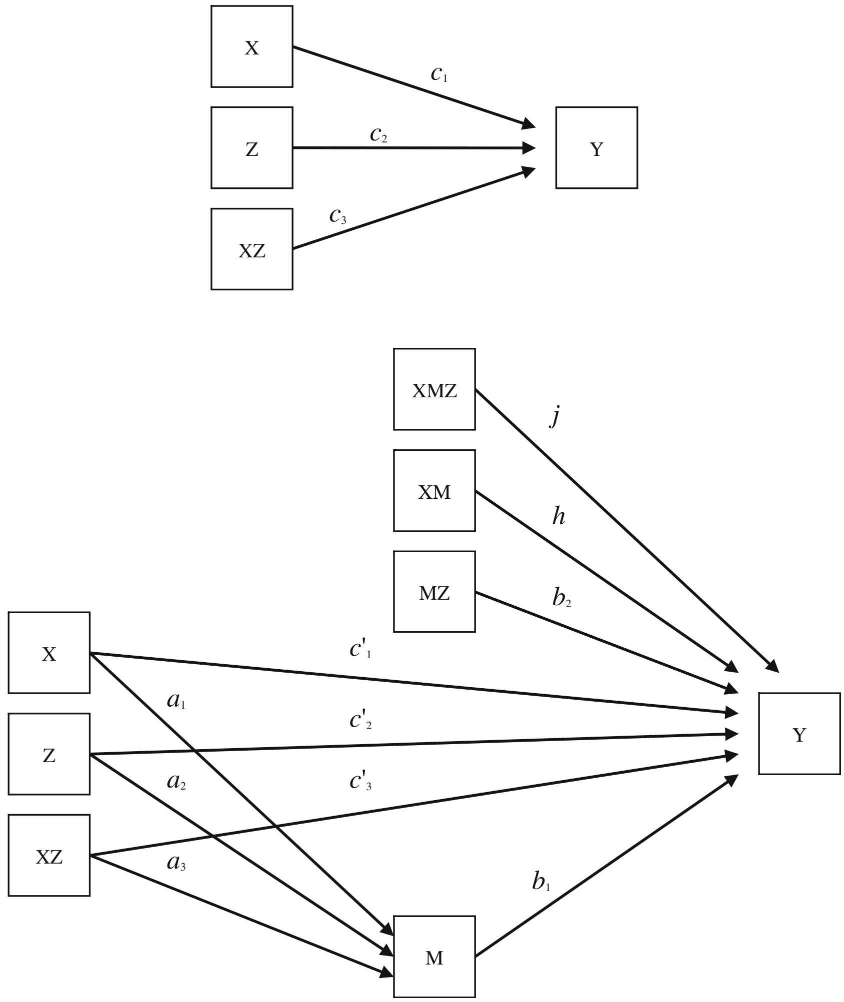

Fig. 6 MacKinnon (2008) General Joint Analysis Mode. *Note.* X= the independent variable, Y= the dependent variable, Z= the moderator variable, M= the mediating variable, XZ= the interaction of X and Z, MZ=the interaction of M and Z, XM= the interaction of X and M, and XMZ= the three-way interaction between X, M, and Z

plots and tests of simple effects to probe the interaction effects. Edwards and Lambert (2007), Preacher et al. (2007), and Tein et al. (2004) provide methods to perform these analyses.

Testing effects: Criteria for the moderation of an indirect effect To examine whether an indirect effect is moderated, it is of interest to investigate whether the mediated effect (*ab*) differs across levels of a fourth, moderating variable. Previous sources have argued that this effect can be defined by either a moderated *a* path, a moderated *b* path, or both moderated *a* and *b* paths in the mediation model (James and Brett 1984; Muller et al. 2005; Preacher et al. 2007; Wegener and Fabrigar 2000), such that if there is moderation in either path of the indirect effect then the mediated relation depends on the level of a moderator variable. There are circumstances, however, in which a heterogeneous *a* or *b* path does not imply a heterogeneous *ab* product term.

Although significant heterogeneity in either the *a* or *b* path may imply significant heterogeneity in the *ab* product term in some cases, examining moderation of the product term or moderation of both paths versus examining moderation of single paths in the mediation model are not conceptually identical. Consider the following example where a moderated *a* path in the mediation model means something different from both moderated *a* and *b* paths in the model. Presume that X is calcium intake, M is bone density, Y is the number of broken bones, and Z is gender. Calcium intake is known to have an effect on the bone density of women, and the relation between calcium intake and bone density is stronger in women than it is in men (i.e., heterogeneity in the *a* path in the model). Specifically, men have greater bone density in general and thus yield fewer gains from supplemental calcium intake. However, bone density affects the fragility of bones in a constant way across males and females, such that low bone density leads to more broken bones (i.e., no heterogeneity of the *b* path in the model). Previous models would deem this scenario as moderation of the indirect effect, arguing that moderation of the *a* path suffices as a test for the effect. There are two problems with this argument. First, testing the heterogeneity of only the a or b path in the mediation model is not a test of mediation because only a single link in the mediated effect is tested in each case. Second, a heterogeneous a path in this model suggests something different from both heterogeneous a and b paths or a heterogeneous ab product. Heterogeneity in both paths of the mediated effect would suggest that gender not only moderates the effect of calcium intake on bone density, but that gender also moderates the effect of bone density on broken bones. Heterogeneity of the product estimate of the mediated effect would suggest that gender moderates the mechanism by which calcium intake affects bone loss; this may or may not be true based on the research literature. Although the moderation of a single path may imply moderation of the product term in some cases, it is critical to differentiate the scenarios as they correspond to different research hypotheses.

There are also numerical examples that show instances when heterogeneity in individual paths of the mediation model does not imply heterogeneity of the product term. Consider the following mediated effect scenarios in two moderator-based subgroups:

	Mediated Effect in Group 1	Mediated Effect in Group 2
Case 1:	(a = -2)(b = -2)	(a = 2)(b = 2)
Case 2:	(a = 1)(b = 2)	(a = 2)(b = 1)

In both scenarios the a and b paths are heterogeneous across groups thus satisfying criteria for the moderation of a mediated effect as defined by Edward and Lambert (2007), James and Brett (1984), Morgan-Lopez and MacKinnon (2006), Preacher et al. (2007), and Wegener and Fabrigar (2000). However the ab product is identical across groups, indicating that there is no moderation of the indirect effect. Although tests for the moderation of a mediated effect based on the heterogeneity of individual path coefficients in the mediation model will be more powerful than a test based on the heterogeneity of the product term (given the usual low power to detect interactions), these tests may also have elevated Type 1 error rates.

Despite potential problems for making inferences on moderation of the ab product from information on the moderation of individual paths, initial simulation work suggests that extending a test of joint significance (where the test for mediation is based on the significance of component paths in the model such that if both $\hat{a}$ and $\hat{b}$ are significant then the mediated effect $\hat{a}\hat{b}$ is deemed significant) to models for mediation and moderation may be acceptable. Specifically, if conclusions about the moderation of $\hat{a}\hat{b}$ are based on whether both $\hat{a}$ and $\hat{b}$ are significantly affected by the moderator variable, Z, Type 1 error rates never exceed .0550 (Fairchild 2008). Effects of the moderator variable on component paths of the mediation model are examined using Eq. 7 and 8, where $\hat{a}_3$ quantifies the effect of Z on the a path and $\hat{b}_2$ quantifies the effect of Z on the b path, respectively (See Fig. 6). If both coefficients are significant, it may be claimed that there is significant moderation of the indirect effect. To obtain either a point estimate or confidence limits for the effect, a product of coefficients test can be used.

To estimate a product of coefficients test for moderation of the indirect effect in the case of a dichotomous moderator variable, separate mediation models can be estimated for each group and equivalence of the $\hat{a}\hat{b}$ point estimates can be compared across moderator-based subgroups. An example of a dichotomous moderator variable might be gender or clinical diagnosis. The null hypothesis associated with the test is that the difference between the two mediated effect point estimates in each group is zero:

$$H_0 : \hat{a}\hat{b}_{group2} - \hat{a}\hat{b}_{group1} = 0 \tag{9}$$

If the point estimates in each group are statistically different from one another, there is significant moderation of the indirect effect (i.e., heterogeneity in the ab product), such that the mediated effect is moderated by group membership. To test the estimate in Eq. 9 for statistical significance, the difference is divided by a standard error for the estimate to form a z statistic. If the groups are independent, the standard error of the difference between the two coefficients is:

$$s_{pooled} = \sqrt{s^2_{\hat{a}\hat{b}_{group1}} + s^2_{\hat{a}\hat{b}_{group2}}} \tag{10}$$

Where $s^2_{\hat{a}\hat{b}_{group1}}$ is the variance of the mediated effect in group 1 and $s^2_{\hat{a}\hat{b}_{group2}}$ is the variance of the mediated effect in group 2.To test heterogeneity of the indirect effect in the case of a continuous moderator variable, variance in the estimates of the ab product across levels of the moderator variable is examined. An example of a continuous moderator variable might be individual motivation to improve. Because a z test as shown above can only accommodate moderator variables with two levels (or a small number of levels with contrasts between two variables) and because levels of continuous moderators often do not represent distinct groups, tests with continuous moderators are more complicated. The question becomes how to assess differences in the ab product across a large number of levels of the moderator variable, and the answer to that question is incomplete at this time. Random coefficient models assess variance in regression coefficients across multiple levels such as multiple levels of a moderator. If the moderator is thought of as a higher order variable across which lower order effects (such as ab) may vary, the random coefficient modeling framework may be suitable to assess variance in the indirect effect across levels of a continuous moderator. Kenny, Korchmaros, and Bolger

(2003) describe an estimate of the variance of *ab* when *a* and *b* are correlated for the case of random effects based on Aroian (1947):

$$s^2_{\left(\hat{a}\hat{b}\right)} = \hat{b}^2 s^2_{\hat{a}} + \hat{a}^2 s^2_{\hat{b}} + s^2_{\hat{a}} s^2_{\hat{b}} + 2\hat{a}\hat{b} s_{\hat{a}\hat{b}} + s^2_{\hat{a}\hat{b}} \quad (11)$$

Bauer, Preacher, and Gil (2006) make an important distinction between the variance of *ab* and the variance of the average value of *ab* in the multilevel model. It is the variance of *ab* that is relevant to the question of whether an indirect effect is moderated as we would like to know if there is substantial variability in *ab* across levels of the moderator variable.

It is also possible to test whether individual paths in the mediation model differ across levels of a moderator variable. These tests can investigate moderation of the direct effect of X on Y (*c'*) or evaluate the generalizability of action and conceptual theory for a program. The null hypothesis to test homogeneity of a program's direct effect is:

$$H_0 : c'_3 = 0 \quad (12)$$

This hypothesis is tested by examining the significance of $\hat{c}'_3$ in Eq. 8; if the regression coefficient is significant, there is significant moderation of the program's direct effect.

Recall that conceptual theory for a program corresponds to the *b* path of the mediation model, which defines the theory that links variables or psychological constructs (e.g., M) to behavioral outcomes. The relationships examined in this piece of the model are driven by previous models or theories presented in the literature that explain motivations for behavior. The null hypothesis to test homogeneity of a program's conceptual theory across levels of the moderator variable is:

$$H_0 : b_2 = 0 \quad (13)$$

This hypothesis is tested by examining the significance of $\hat{b}_2$ in Eq. 8; if the regression coefficient is significant there is significant moderation of the program conceptual theory.

Action theory for a program corresponds to the *a* path of the mediation model, which defines what components of the program are designed to manipulate the mechanisms of change. This piece of the model illustrates how the program intervenes to modify hypothesized mediators. The null hypothesis to test homogeneity of a program's action theory across levels of the moderator variable is:

$$H_0 : a_3 = 0 \quad (14)$$

This hypothesis is tested by examining the significance of $\hat{a}_3$ in Eq. 7; if the regression coefficient is significant there is significant moderation of the program action theory. See Chen (1990) for more details on program action and conceptual theory.

Testing effects: Criteria for the mediation of a moderator effect A test for the mediation of a moderator effect examines whether the magnitude of an overall interaction effect of the independent variable (X) and the moderator variable (Z) on the dependent variable (Y) is reduced once the mediator is accounted for in the model (Muller et al. 2005). Thus, the examination of the mediation of a moderator effect considers the mediation model as a means to explain why a treatment effect of X on Y is moderated by a third variable, Z. In this way, the mediation of a moderator effect hypothesis probes mediation as a possible process that accounts for the interaction of the treatment and the outcome. There is no need to differentiate methods for categorical and continuous moderator variables here.

One way to test whether an XZ interaction in the data is accounted for, at least in part, by a mediating relation is to examine whether the magnitude of the regression coefficient corresponding to the overall interaction effect, c_3, is reduced once the mediator is added to the model. Using coefficients from Equations 6 and 8, where $\hat{c}'_3$ represents the direct effect of the interaction effect on Y once the mediator is included in the model, a point estimate and standard error of this difference can be computed:

$$\hat{c}_3 - \hat{c}'_3 \quad (15)$$

$$s_{\hat{c}_3 - \hat{c}'_3} = \sqrt{s^2_{\hat{c}_3} + s^2_{\hat{c}'_3} - 2 s_{\hat{c}_3 \hat{c}'_3}} \quad (16)$$

Dividing an estimate of the difference in Eq. 15 by its standard error in Eq. 16 provides a test of significance for the estimate. If there is a significant difference in the coefficients then the overall moderation effect of XZ on Y is significantly explained, at least in part, by a mediated relation. The null hypothesis corresponding to this effect is that there is no reduction in the overall interaction once accounting for the mediator:

$$H_0 : c_3 - c'_3 = 0 \quad (17)$$

A test of the mediation of a moderator effect with a product of coefficients estimator circumvents the need for testing the overall interaction of XZ on Y as shown above. Morgan-Lopez and MacKinnon (2006) presented a point estimator and standard error for the product of coefficients method:

$$\hat{a}_3 \hat{b}_1 \quad (18)$$

$$s_{\hat{a}_3 \hat{b}_1} = \sqrt{s^2_{\hat{a}_3} \hat{b}^2_1 + s^2_{\hat{b}_1} \hat{a}_3{}^2} \quad (19)$$

The product of coefficients estimator in Eq. 18 illustrates that the effect of the interaction of the independent and moderator variables on the outcome (represented by $\hat{a}_3$) is transmitted through a mediating variable (represented by $\hat{b}_1$). Dividing the estimate of the product by an estimate of its standard error in Eq. 19 provides a test of significance for the estimate based on normal theory. However, as described for the product of coefficients estimator in the single mediator model, asymmetric confidence intervals for the distribution of the product are more accurate than tests based on the normal distribution and should be implemented (note also that a test of joint significance of the coefficients could be conducted). If the product is significantly different from zero, then the moderator effect is explained, at least in part, by a mediating mechanism. The null hypothesis corresponding to this test is that the product of the coefficients is equal to zero:

$$H_0 : a_3 b_1 = 0 \tag{20}$$

Numerical Examples

Returning now to the research scenarios presented at the beginning of this paper, numerical examples can be provided to show how these analyses may appear in practice. Two simulated datasets were created to explore the questions. Scenario #1 asked "*Is the process by which a program has an effect the same across different types of participants?*" The researcher investigates whether the indirect effect is moderated to answer this question. Note that although it is necessary to estimate the XM interaction in Equation 8 to avoid bias in the XZ term (Yzerbyt, Muller, and Judd 2004), the XMZ interaction need not be estimated if there is no supporting hypothesis for its existence. Likewise, Equation 6 need not be estimated if overall effects are not part of the research question. Recall that for this example the independent variable, X, was a worksite-wellness program to reduce obesity-related health risk. The mediating variable, M, was knowledge of the benefits of eating fruits and vegetables, and the outcome, Y, was employee consumption of fruits and vegetables. The moderator variable in analysis, Z, was family history of obesity-related illness with 107 employees having no history of disease (i.e., $n_{group1} = 107$) and 93 employees having a history of disease (i.e., $n_{group2} = 93$). The null hypothesis for the research question in Scenario #1 is that there is no difference in the indirect effect of the program on employee fruit and vegetable consumption through knowledge of fruit and vegetable benefits across employees who have a family history of disease versus those who have no family history of disease. To test this hypothesis, an estimate of the mediated effect can be computed in each group and compared for differences. Estimating Equations 2 and 3, the mediated effect, $\hat{a}\hat{b}\left(s_{\hat{a}\hat{b}}\right)$, for employees with no family history of obesity-related disease was computed as .0004(.0186), with an asymmetric confidence interval of − .0400, .0413, where $\hat{a}\left(s_{\hat{a}}\right) = .0041(.2018)$and $\hat{b}\left(s_{\hat{b}}\right) = .0923(.0969)$ were used as inputs into the PRODCLIN program (MacKinnon et al. 2007). The mediated effect for employees with a family history of obesity-related disease was .2158(.0821), with an asymmetric confidence interval of .0693, .3909, where $\hat{a}\left(s_{\hat{a}}\right) = .5479(.1854)$ and $\hat{b}\left(s_{\hat{b}}\right) = .3938(.0687)$ were used as inputs into the PRODCLIN program (MacKinnon et al. 2007). Initial inspection of the indirect effects shows that the mediation relation in the data is significant for individuals with a family history of obesity-related illness (the simple mediated effect for group 1) and nonsignificant for individuals with no family history of disease (the simple mediated effect for group 2). Statistical inspection of the difference between the two indirect effects shows a difference of .2154. Using Eq. 11, the pooled standard error for this difference was computed as .0421, yielding a *z*-statistic of 5.1143. Because the obtained *z*-value is greater than the critical value (1.96) associated with α = .05, there is significant moderation in the indirect effect such that *ab* is heterogeneous across moderator-based subgroups. Alternatively estimating Equations 7 and 8 and examining a test of joint significance of the $\hat{a}_3\left(s_{\hat{a}_3}\right)$ and $\hat{b}_2\left(s_{\hat{b}_2}\right)$ coefficients to examine whether the indirect effect is moderated yields the same conclusion. Both parameter estimates are significant: $\hat{a}_3\left(s_{\hat{a}_3}\right) = .2760(.1386)$ and $\hat{b}_2\left(s_{\hat{b}_2}\right) = .1919(.0697)$, respectively. Because the moderator variable in this example only had two levels, there is no need to follow up the finding with simple effects; the difference in the mediated effects occurs between the simple mediated effect in group 1 and the simple mediated effect in group 2. Had there been three or more levels of the moderator variable in the example, the researcher could conduct follow up analyses to investigate simple mediated effects at different levels of the moderator variable to see where specific differences occur. Had there also been a significant XM interaction in the data, the researcher would need to estimate simple mediated effects across levels of X as well.

Scenario #2 asked "*Can a mediation relation explain an interaction effect in my data?*" To answer this question, the researcher investigates whether the interaction effect can predict a mediating variable which in turn predicts the outcome. Recall that X was the worksite-wellness program described above, M was a social norm for eating fruits and vegetables, Z was part-time or full-time work status (191 part-time employees, 209 full-time employees), and Y was fruit and vegetable consumption. The null hypothesis for the research question is that a mediating variable cannot explain any part of the interaction. To test this hypothesis, the $\hat{a}_3\hat{b}_1$ point estimate is computed for the data and tested for significance. As indicated in the numerical example for Research Scenario #1, although it is necessary to estimate the XM interaction to avoid bias in the XZ term (Yzerbyt et al. 2004), there is no need to model the XMZ interaction if there is no hypothesis to support its estimation. Estimating Eq. 7 and 8, $\hat{a}_3\hat{b}_1$ was computed as .1781 (.0814), with an asymmetric confidence interval of .0192, .3383, where $\hat{a}_3\left(s_{\hat{a}_3}\right) = .1550(.0706)$ and $\hat{b}_1\left(s_{\hat{b}_1}\right) = 1.1495(.0405)$ were used as inputs into the PRODCLIN program (MacKinnon et al. 2007). Failure to include zero in the confidence interval indicates significance of the point estimate, such that the interaction of the program effect and part-time versus full-time work status is explained at least in part by the mediating mechanism of social norms for eating fruits and vegetables. Example SAS datasets, program code and output, and complete worked examples for the general model to simultaneously test mediation and moderation effects are available from the first author.

Issues of Power in Models to Test Mediation and Moderation Effects

Power to detect interaction effects is often low because of the small effect sizes observed in social science (Aiken and West 1991). Models that simultaneously examine mediation and moderation effects are at an even greater disadvantage as they involve several interaction terms as well as estimation of indirect effects. Fairchild (2008) investigated power for the general model using previous arguments made in the literature that interaction effects in real data typically range from explaining 1% to 3% of the variance in the dependent variable (Champoux and Peters 1987; Evans 1985; McClelland and Judd 1993). Results showed that when the effect size of the $\hat{a}_3$ and $\hat{b}_2$ parameters both explained 1% of the variance, a test of joint significance for moderation of the indirect effect required $N \geq 1000$ for .8 power. If the effect size of the parameters both explained 3% of the variance, the sample size requirement for .8 power reduced to $N \geq 500$. Morgan-Lopez and MacKinnon (2006) showed similar results for tests to examine the mediation of an interaction effect.

The large sample size requirement of these tests makes investigation of mediation and moderation effects in small samples challenging. Effect size analysis usually offers a guide to the practical utility of effects that is independent of significance and sample size, but effect size measures for models that simultaneously analyze mediation and moderation effects are not yet fully researched. Although increasing sample size has historically been most utilized to increase the power of a test, there are several alternatives that may increase power when accessing additional data is not possible (Beck 1994).

Statistical significance of a test is based on four parameters: sample size, effect size, power, and the alpha-level set for the experiment. Thus to achieve .8 power when sample size is fixed either the effect size of the treatment and/or the alpha level of the experiment may be manipulated to increase sensitivity of the research design. For example, rather than implementing the nominal .05 criterion that allows for errors in null-hypothesis-significance-testing 5 times out of every 100, the researcher may choose to evaluate results against a more liberal .10 criterion that allows for errors 10% of the time. Increasing the Type 1 error rate for the study decreases the possibility of making a Type 2 error, or the failure to find a significant effect when one truly exists in the data. This option may be appropriate when an effect of interest is especially important and avoiding Type 2 errors is critical. A second option for increasing power when sample size is limited in a study is to increase the effect size of the manipulation. This option involves increasing the between-group variance on the independent variable in the design and decreasing the within-group variance, or error variance, of the variables (Beck 1994). To increase the between-group variance on the independent variable the researcher may impose an extreme-groups design (see Borich and Godbout 1974) or deliver a stronger manipulation so that differences between the treatment and control groups are exaggerated. When the strength of the intervention is not easily increased or formation of extreme groups is not feasible, researchers may opt to focus on decreasing within-group variance, or error variance, in the design. Statistical methods to reduce error variance include improving reliability of the measurement tools used in analysis and implementing analysis of covariance models that partition covariates out of treatment effects. Research design tools that are able to reduce error variance in a model include blocking and matching designs that attempt to equate treatment and control groups on non-intervention dimensions in which they may differ (Cook and Campbell 1979).

Recommendations

To best implement the simultaneous analysis of mediation and moderator effects, the researcher should first identify which third variable model characterizes his or her primary research question. This a priori model identification acts as the starting point of interpretation and can be guided by theory or previous research. If the primary research question is determined to be mediation, the researcher may examine all possible interaction effects in the model or a subset of theoretically-relevant interactions and discuss which moderated effects exist.

Methods for examining the moderation of the indirect effect have been discussed in this manuscript, as well as methods for examining moderation in individual links of the mediation model such as the direct effect. Exploring the moderation of either the total or direct effect in the mediation model is simply a means of examining the significance of the individual regression coefficients corresponding to those effects (i.e., c_3 and c'_3, respectively), and any significant interaction effect in the model will be interpreted with reference to mediation. Interactions may be further probed with simple slopes analyses to examine variable relations across levels of the moderator variable (Edwards and Lambert 2007; Preacher et al. 2007; Tein et al. 2004).

If the statistical model appropriate to answer the primary research question is determined to be moderation, mediation may be investigated as a means to explain an overall moderated treatment effect. To that end, the interpretation of interaction effects in the model focus on whether a mediating mechanism is responsible, at least in part, for overall moderation in the data. As presented in the manuscript, this assessment may be made with either a difference in coefficients ($\hat{c}_3 - \hat{c}'_3$) or product of coefficients ($\hat{a}_3\hat{b}_1$) estimator. Both tests examine whether there is a decrease in the overall XZ interaction once a mediating variable is modeled. However, given the availability of more accurate significance testing for the product of coefficients estimator, it is recommended that researchers analyze $\hat{a}_3\hat{b}_1$. Preferably, the exploration of possible interactions in either model are well-thought-out and delineated by theory to avoid confusion and unwieldy models, as well as to facilitate interpretation of effects in the model. Regardless of which third variable model is identified with the primary research question, it is necessary to model the XM interaction to avoid bias in the XZ product term (Yzerbyt, Muller, and Judd 2004).

Limitations and Future Directions

The cost of the generalizability of the general model to test mediation and moderation effects is possible inflation of Type I error, lack of power, and difficulty with interpretation of model parameters if several effects are present. The model may be simplified, however, to represent more specialized cases of mediation and moderation joint effects such as baseline by treatment interactions by constraining paths in the model to be zero.

An additional limitation of models with moderation and mediation is the extensive assumptions required for accurate assessment of relations among variables (Holland 1988). The sensitivity of conclusions to violations of assumptions is not yet known and correct conclusions will likely require repeated applications in any substantive research area. In particular, often the X variable is the only variable that represents random assignment, making interpretation of causal relations between other variables in the model susceptible to omitted variable bias. In many applications, the model results may represent descriptive information about how variables are related rather than elucidating true causal relations among variables. Information on true causal relations will require programs of research to replicate and extend results as well as information from other sources such as qualitative information and replication studies in different substantive research areas.

Estimation of the methods and their performance can be investigated for categorical and longitudinal data. Previous work has explored the needs of categorical data in the single mediator model (MacKinnon and Dwyer 1993) and this research may serve as a resource for extending the general model to test mediation and moderation effects into that domain. With regard to longitudinal data, the current model can be applied to two-wave data by using difference scores or residualized change scores, but data with three or more waves cannot yet be accommodated. The addition of cross-lagged effects in longitudinal data frameworks such as autoregressive models or latent growth models increases the possible number of interactions among variables and whether the investigation of all possible cross-lagged effects is valuable or overly cumbersome will need to be determined.

In summary, many questions in prevention science involve how and for whom a program achieves its effects. Mediation and moderation models are ideal for examining these questions. By investigating both mediation and moderation effects in data from prevention programs information about the mechanisms underlying program effects as well as the generalizability of program effects and curricula can be evaluated. Models that simultaneously estimate mediation and moderation effects not only allow for the examination of these questions, but also permit the evaluation of more complex research hypotheses such as whether a moderator effect in the data can be explained by a mediating mechanism, or whether a mediating mechanism depends on the level of another variable. A broad integration of these models into the substantive research literature will enhance the information drawn from prevention work and will inform our knowledge on prevention models.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Aroian, L. A. (1947). The probability function of the product of two normally distributed variables. *Annals of Mathematical Statistics*, *18*, 265–271. doi:10.1214/aoms/1177730442.

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. doi:10.1037/0022-3514.51.6.1173.

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.

Beck, C.T. (1994). Achieving statistical power through research design sensitivity. *Journal of Advanced Nursing*, *20*, 912–916.

Borich, G. D., & Godbout, R. C. (1974). Extreme groups designs and the evaluation of statistical power. *Educational and Psychological Measurement*, *34*, 663–675.

Champoux, J. E. & Peters, W. S. (1987). Form, effect size, and power in moderated regression analysis. *Journal of Occupational Psychology*, *60*, 243–255.

Chen, H. T. (1990). *Theory driven evaluations*. Newbury Park, CA: Sage.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. A. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.

Dearing, E., & Hamilton, L. C. (2006). Contemporary advances and classic advice for analyzing mediating and moderating variables. *Monographs of the Society for Research in Child Development*, (December), 71, 88–104.

Donaldson, S. I. (2001). Mediator and moderator analysis in program development. In S. Sussman (Ed.), *Handbook of program development for health and behavior research and practice* (pp. 470–496). Thousand Oaks, CA: Sage.

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. doi:10.1037/1082-989X.12.1.1.

Evans, M. G. (1985). A Monte Carlo study of the effects of correlated method variance in moderated multiple regression analysis. *Organizational Behavior and Human Decisions Processes*, *36*, 305–323.

Fairchild, A. J. (2008). *A comparison of frameworks for the joint analysis of mediation and moderation effects*. Unpublished doctoral dissertation. Arizona State University, Tempe.

Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, *51*, 115–134. doi:10.1037/0022-0167.51.1.115.

Gogineni, A., Alsup, R., & Gillespie, D. (1995). Mediation and moderation in social work research. *Social Work Research*, *19*, 57–63.

Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equations models. *Sociological Methodology*, *18*, 449–484. doi:10.2307/271055.

Hoyle, R. H., & Robinson, J. C. (2003). Mediated and moderated effects in social psychological research: Measurement, design, and analysis issues. In C. Sansone, C. Morf, & A. T. Panter (Eds.), *Handbook of methods in social psychology* (pp. 213–233). Thousand Oaks, CA: Sage.

James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *The Journal of Applied Psychology*, *69*, 307–321. doi:10.1037/0021-9010.69.2.307.

Judd, C. M., & Kenny, D. A. (1981a). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, *5*, 602–619. doi:10.1177/0193841X8100500502.

Judd, C. M., & Kenny, D. A. (1981b). *Estimating the effects of social intervention*. Cambridge, England: Cambridge University Press.

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

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

MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*, *17*, 144–158. doi:10.1177/0193841X9301700202.

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.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science*, *1*, 173–181. doi:10.1023/A:1026595011371.

MacKinnon, D. P., & Lockwood, C. M. (2001). *Distribution of product tests for the mediated effect: Power and Type I error rates*. Unpublished manuscript.

MacKinnon, D. P., Lockwood, C. M., & Hoffman, J. M. (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 the significance of mediation and other intervening variable effects. *Psychological Methods*, *7*, 83–104. doi:10.1037/1082-989X.7.1.83.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, *39*, 99–128. doi:10.1207/s15327906mbr3901_4.

McClelland, G. H. & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, *114*, 376–390.

Morgan-Lopez, A. A., & MacKinnon, D. P. (2006). Demonstration and evaluation of a method for assessing mediated moderation. *Behavior Research Methods*, *38*, 77–87.

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology*, *89*, 852–863. doi:10.1037/0022-3514.89.6.852.

Preacher, K. J., Rucker, D. D., & Hayes, R. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*, 185–227.

Rose, B. M., Holmbeck, G. N., Coakley, R. M., & Franks, E. A. (2004). Mediator and moderator effects in developmental and behavioral pediatric research. *Journal of Developmental and Behavioral Pediatrics*, *25*, 58–67. doi:10.1097/00004703-200402000-00013.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290–312. doi:10.2307/270723.

Sobel, M. E. (1986). Some new results on indirect effects and their standard errors in covariance structure models. *Sociological Methodology*, *16*, 159–186. doi:10.2307/270922.

Tein, J., Sandler, I. N., MacKinnon, D. P., & Wolchik, S. A. (2004). How did it work? Who did it work for? Mediation in the context of a moderated prevention effect for children of divorce. *Journal of Consulting and Clinical Psychology*, *72*, 617–624. doi:10.1037/0022-006X.72.4.617.

Wegener, D. T., & Fabrigar, L. R. (2000). Analysis and design for nonexperimental data addressing causal and noncausal hypotheses. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 412–450). New York: Cambridge University Press.

West, S. G., & Aiken, L. S. (1997). Toward understanding individual effects in multiple component prevention programs: Design and analysis strategies. In K. Bryant, M. Windle, & S. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 167–209). Washington, DC: American Psychological Association.

Yzerbyt, V. Y., Muller, D., & Judd, C. M. (2004). Adjusting researchers' approach to adjustment: On the use of covariates when testing interactions. *Journal of Experimental Social Psychology*, *40*, 424–431. doi:10.1016/j.jesp.2003.10.001.