A graphical representation of the mediated effect

MATTHEW S. FRITZ
Virginia Polytechnic Institute and State University, Blacksburg, Virginia

AND

DAVID P. MACKINNON
Arizona State University, Tempe, Arizona

Mediation analysis is widely used in the social sciences. Despite the popularity of mediation models, few researchers have used graphical methods, other than structural path diagrams, to represent their models. Plots of the mediated effect can help a researcher better understand the results of the analysis and convey these results to others. This article presents a method for creating and interpreting plots of the mediated effect for a variety of mediation models, including models with (1) a dichotomous independent variable, (2) a continuous independent variable, and (3) an interaction between an independent variable and the mediating variable. An empirical example is then presented to illustrate these plots. Sample code for creating plots of the mediated effect in R and SAS is also included, and may be downloaded from www.psychonomic.org/archive.

Graphical representations of data are often used by researchers to explore relations among variables. Plots and graphs can be used for initial data exploration, to examine violations of model assumptions, and to show predicted scores from a final model. Despite the versatility and usefulness of plots, researchers have rarely used them to illustrate the results from mediation studies. Instead, researchers present path diagrams that show the structural relationship between the independent, mediator, and dependent variables. However, plots that not only show the structural properties of the mediation model, but actually illustrate the magnitude of the mediated effect itself, can be very useful as a visual representation of the mediated effect, which may be more desirable than presenting numerical results alone. These plots may be especially useful when investigating complex mediation models such as inconsistent mediation models in which the direct effect and mediated effect have opposite signs, or models in which the mediator interacts with the independent variable, which were first described by Merrill (1994). MacKinnon (2008) used a similar approach to describe more general plots of the effects in mediation models.

The purpose of this article is to explain how to construct plots of the mediated effect for a study with a dichotomous (e.g., random assignment to a treatment and control group) or a continuous independent variable, a continuous mediator, and a continuous outcome variable. An empirical example is then used to show what the finished plots look like and how to interpret plots of the mediated effect. Finally, sample code for creating plots of the mediated effect is given for two common statistical software packages, SAS (SAS Institute, 2006) and R (R Core Development Team, 2006), to assist researchers in making their own plots. The goal of this article is to give researchers another tool for presenting and explaining results from mediational studies that may be used in presentations and publications.

Mediation

Mediation is the name given to models in which the effect of an antecedent or independent variable (X) on a dependent variable (Y) is transmitted through a third intervening or mediating variable (M). In other words, X affects M, which in turn affects Y. Figure 1 shows a diagram of a single mediator model, where the top diagram represents the total effect of X on Y and the bottom diagram represents the mediated effect of X on Y through M. These two path diagrams can be represented using the regression equations

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

$$M = i_2 + aX + e_2 \tag{2}$$

$$Y = i_3 + c'X + bM + e_3, \tag{3}$$

where c is the total effect of X on Y, c' is the effect of X on Y adjusted for M, b is the effect of M on Y adjusted for X, a is the effect of X on M, i_1, i_2, and i_3 are the intercepts, and e_1, e_2, and e_3 are the residuals. The a path is known as the *action theory* because X is often directly manipulated, whereas the b path is known as the *conceptual theory* (Chen, 1990; MacKinnon, Taborga, & Morgan-Lopez, 2002), because M is not usually directly manipulated.

The mediated or indirect effect is equal to ab or the product of the action theory and the conceptual theory coefficients. The effect of X on Y that does not pass through M is called the *direct effect* and is equal to c'. Models where $c' = 0$ are called *completely mediated models*, and models where $c' \neq 0$ are called *partially mediated*
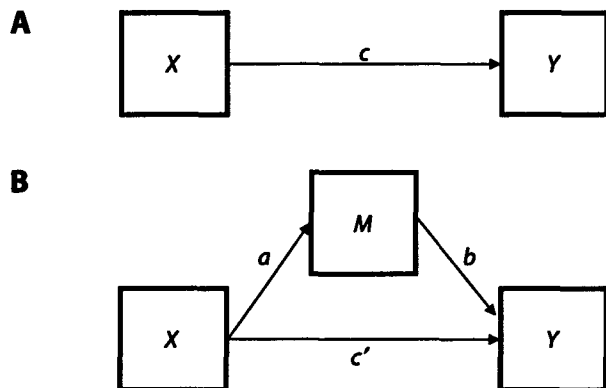
M. S. Fritz, matt.fritz@vt.edu

A

B

Figure 1. Path diagrams for (A) the total effect of the independent variable on the dependent variable and (B) the indirect effect of the independent variable on the dependent variable through the mediator variable.

models (Baron & Kenny, 1986). In ordinary-least-squares regression,

$$c = ab + c', \qquad (4)$$

although this relationship does not hold for all logistic regression or multilevel models (MacKinnon, 2008). Mediation should not be confused with an interaction, also known as moderation, in which the magnitude of the effect of X on Y is determined by the value of M; in a completely mediated model, if M is held constant, then X has no effect on Y.

There are many statistical tests of mediation, and these tests fall into three categories: causal steps, product of coefficients, and difference in coefficients. In the causal steps tests, each step in the causal chain is tested for significance in sequence. For example, the Baron and Kenny (1986) causal steps test first tests the significance of the effect of X

on Y, c, then the significance of the effect of X on M, a, and then finally the effect of M on Y controlling for X, b; if all three effects are significant, then mediation is present. The product of coefficients tests directly test the indirect effect, ab, for significance, by dividing ab by its standard error and comparing the resulting value with a normal distribution, or by creating confidence intervals around the indirect effect, as in MacKinnon, Fritz, Williams, and Lockwood (2007). Finally, the difference in coefficients tests investigate the difference between the overall effect and the part of the overall effect that does not pass through the mediator, $c - c'$, by dividing the difference by its standard error and comparing the value with a t distribution for significance. For more information on tests of mediation, see MacKinnon, Lockwood, Hoffman, West, and Sheets (2002) and MacKinnon, Lockwood, and Williams (2004).

Despite the variety of different statistical tests of mediation, all of the tests operate under the same assumptions. The first assumption is that the order of causation is correct—that is, X causes M, even when X is not random assignment to treatment, and M causes Y. Also, the interaction between X and M is assumed to be zero in Equation 3, although this assumption can be relaxed, as described later in this article. Other assumptions are that the dependent variables have a normal distribution and the residuals are independent and normally distributed. These assumptions apply to all of the plots.

Plotting the Mediated Effect

To plot a mediated effect for a dichotomous independent variable, Equations 1, 2, and 3 must first be estimated to find values for the intercepts and regression coefficients (MacKinnon, 2008; Merrill, 1994). Note that X was dummy coded 0 and 1, and for convenience, hats were not included over the parameter estimates in the plots. Second,
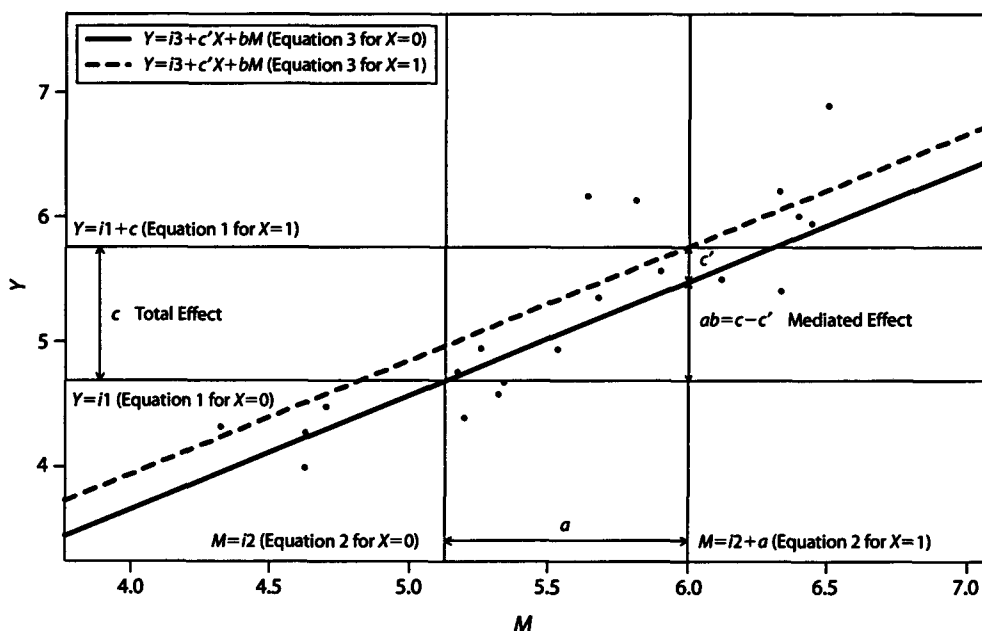

Figure 2. Plot of the mediated effect for a dichotomous X variable.

the data are plotted with Y on the vertical axis and M on the horizontal axis, as shown in Figure 2. Next, Equation 1 is plotted for the values of X (i.e., 0 and 1), such that a horizontal line is placed at $Y = \hat{i}_1$, corresponding to $X = 0$, and a second horizontal line is placed at $Y = \hat{i}_1 + \hat{c}$, corresponding to $X = 1$. The distance between the horizontal lines represents the total effect of X on Y, $\hat{c}$. Then, Equation 2 is plotted for both values of X, resulting in a vertical line at $M = \hat{i}_2$ and a second vertical line at $M = \hat{i}_2 + \hat{a}$. The distance between the two vertical lines represents the action theory and is equal to $\hat{a}$. Finally, Equation 3 is plotted at both values of X. The slopes of these parallel simple regression lines are equal to $\hat{b}$. Looking at the $M = \hat{i}_2 + \hat{a}$ vertical line, the distance between the $Y = \hat{i}_1 + \hat{c}$ horizontal line and the point where the regression line for the $X = 0$ group intersects the $M = \hat{i}_2 + \hat{a}$ vertical line is equal to $\hat{c}'$, and the distance from this intersection point to the $Y = \hat{i}_1$ horizontal line is equal to the mediated effect, $\hat{a}\hat{b}$.

Interpreting the plot is straightforward (MacKinnon, 2008; MacKinnon, Fairchild, & Fritz, 2007; Merrill, 1994). Figure 2 shows that $\hat{c}$ is the predicted amount of change in Y for a one unit change in X and that the larger the distance between the horizontal lines, the larger the overall effect of X on Y. The same is true of $\hat{a}$, which is the predicted amount of change in M for a one unit change in X; the larger the distance between the vertical lines, the greater the effect of X on M. The slope of the simple regression lines, $\hat{b}$, is then the predicted change in Y for a one unit change in M, adjusted for X. Regardless of whether the effect is positive or negative, the steeper the slope of the simple regression lines, the larger the effect of M on Y, adjusting for X. The mediated effect, $\hat{a}\hat{b}$, is then the predicted change in Y for an $\hat{a}$ unit increase in M, which corresponds to a one unit change in X. Finally, looking at the length of $\hat{a}\hat{b}$ relative to

the length of $\hat{c}$ indicates the amount of the overall effect of X on Y that is mediated by M; in this example, the mediator accounts for about 60% of the effect of X on Y. Regardless of the relative size of $\hat{a}\hat{b}$, statistical significance of the mediated effect should be investigated using one of the tests of mediation described earlier.

For a continuous independent variable, Equations 1, 2, and 3 are again estimated to find values for the intercepts and regression coefficients, and the data are plotted with Y on the vertical axis and X on the horizontal axis, as shown in Figure 3. Next, Equations 1 and 2 are plotted for the mean of X and one unit above the mean of X, such that horizontal lines are plotted at $Y = \hat{i}_1 + \hat{c}\bar{X}$, which is equal to the mean of Y, and at $Y = \hat{i}_1 + \hat{c}(\bar{X} + 1)$, and vertical lines are plotted at $M = \hat{i}_2 + \hat{a}\bar{X}$, which equals the mean of M, and $M = \hat{i}_2 + \hat{a}(\bar{X} + 1)$. The distance between the horizontal lines is then equal to $\hat{c}$, and the distance between the two vertical lines is equal to $\hat{a}$. Finally, Equation 3 is plotted. Because X is continuous, specific values of X must be selected to plot. In Figure 3, the values of X selected are the mean of X and one standard deviation above the mean of X, although any values of X may be selected and any number of values may be plotted (Aiken & West, 1991); the slope of each of these lines is equal to $\hat{b}$. Looking at the $M = \hat{i}_2 + \hat{a}(\bar{X} + 1)$ vertical line, the distance between the $Y = \hat{i}_1 + \hat{c}(\bar{X} + 1)$ horizontal line and the point where the regression line for the mean of X intersects the $M = \hat{i}_2 + \hat{a}(\bar{X} + 1)$ vertical line is equal to $\hat{c}'$, and the distance from the intersection point to the $Y = \hat{i}_1 + \hat{c}\bar{X}$ horizontal line is equal to the mediated effect, $\hat{a}\hat{b}$. The interpretation of the effects in the plot are identical to the case with a dichotomous X variable.

In both Figure 2 and Figure 3, all of the estimates of the effects in the model, $\hat{a}$, $\hat{b}$, $\hat{c}$, and $\hat{c}'$, were positive, but this
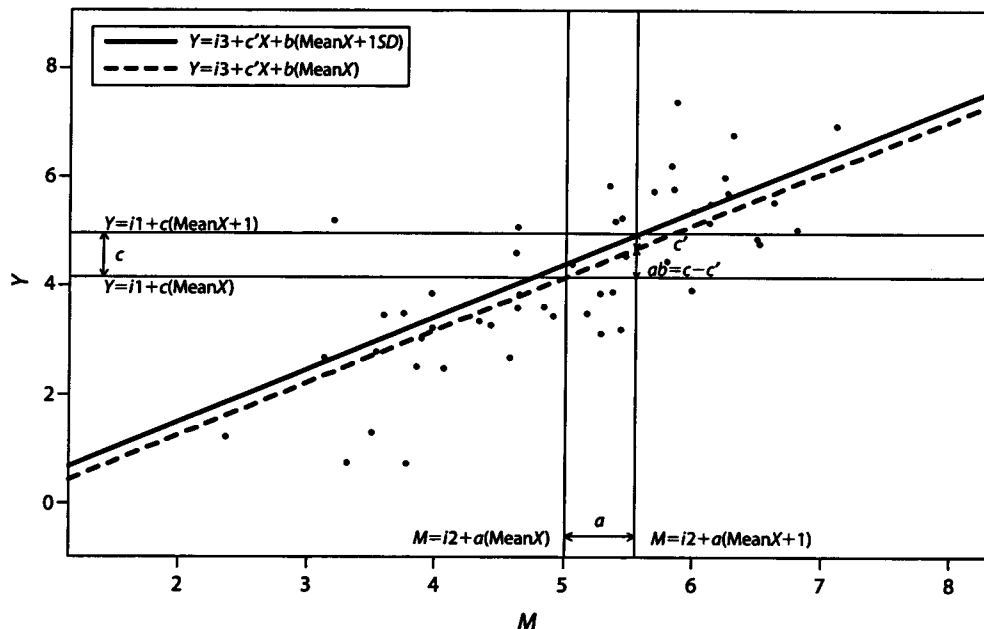

Figure 3. Plot of the mediated effect for a continuous X variable. Note that SD stands for the standard deviation of X.

will not necessarily occur when plotting real data. When $\hat{b}$ is negative, the simple regression lines slope downward. When $\hat{c}$ is negative, the distance between the horizontal lines will remain the same, but the order of the lines will be reversed and $\hat{a}\hat{b}$ and $\hat{c}'$ will switch places. When $\hat{a}$ is negative, the distance between the vertical lines will remain the same, but the order of the lines will reverse, although $\hat{a}\hat{b}$ and $\hat{c}'$ will still occur on the $M = \hat{i}_2 + \hat{a}$ line. For instance, in Figure 2, if $\hat{a}$ was negative, the line $M = \hat{i}_2 + \hat{a}$ would be $\hat{a}$ units to the left of the $M = \hat{i}_2$ line.

A word of caution is needed for inconsistent mediation models. In an inconsistent mediation model, $\hat{a}\hat{b}$ and $\hat{c}'$ are significant, but opposite in sign, which can cause $\hat{c}$ to be nonsignificant. As shown in Equation 4, if $\hat{a}\hat{b}$ and $\hat{c}'$ are equal in magnitude but opposite in sign, then $\hat{c}$ would be equal to 0. When this occurs, the slopes of the simple regression lines may be positive or negative and the vertical lines may or may not be switched, depending on the signs of the $\hat{a}$ and $\hat{b}$ coefficients. It is also possible in inconsistent mediation models for the mediated effect to be larger than the total effect, such that the simple regression slopes of Equation 3 cross the $M = \hat{i}_2 + \hat{a}$ line outside of the range of the horizontal lines that represent the total effect $\hat{c}$.

XM Interactions

In certain cases, the effect of the mediator is not equal across values of X. In these cases, an interaction between the independent variable and the mediating variable, known as an XM interaction, is present. A significant XM interaction indicates that the effect of M on Y is a function of X, and changes as the value of X changes. Plotting a mediation model with an XM interaction is very similar to plotting the typical mediation model (Fritz & MacKinnon, 2006; Merrill, 1994). First, Equations 1 and 2 are

estimated to find values for the coefficients. Rather than estimating Equation 3, however, the equation

$$Y = i_3 + c'X + bM + gXM + e_3 \qquad (5)$$

is estimated. The only difference between Equation 3 and Equation 5 is the inclusion of the XM interaction in Equation 5.

Plotting of the mediated effect with an XM interaction then directly follows the steps used to plot mediation models without an interaction; an example for a dichotomous independent variable is shown in Figure 4. In Figure 4, the solid line represents the simple regression line for the group coded $X = 0$ and the dashed line represents the simple regression line for the group coded $X = 1$. The dotted line represents the simple regression line for the group coded $X = 1$ when the interaction effect $\hat{g}$ is set to 0 and the distance between the dashed and dotted lines at a specific value of M is equal to the interaction effect at that value of M, $\hat{g}M$.

Empirical Example

To better illustrate the construction and use of these plots, an empirical example is included using data from the Athletes Training and Learning to Avoid Steroids (ATLAS) program (Goldberg et al., 1996; Goldberg et al., 2000). The ATLAS program was designed for high school football players, with the goal of decreasing their intentions to use anabolic steroids by offering improved dietary behavior and increased strength training self-efficacy as direct alternatives to steroid use. The diet and strength-training methods produced similar results to steroid use without the negative consequences associated with using illicit drugs. ATLAS changed these three outcome variables (intentions to use steroids, nutrition behavior, and strength training self-efficacy) by targeting 12 main me-
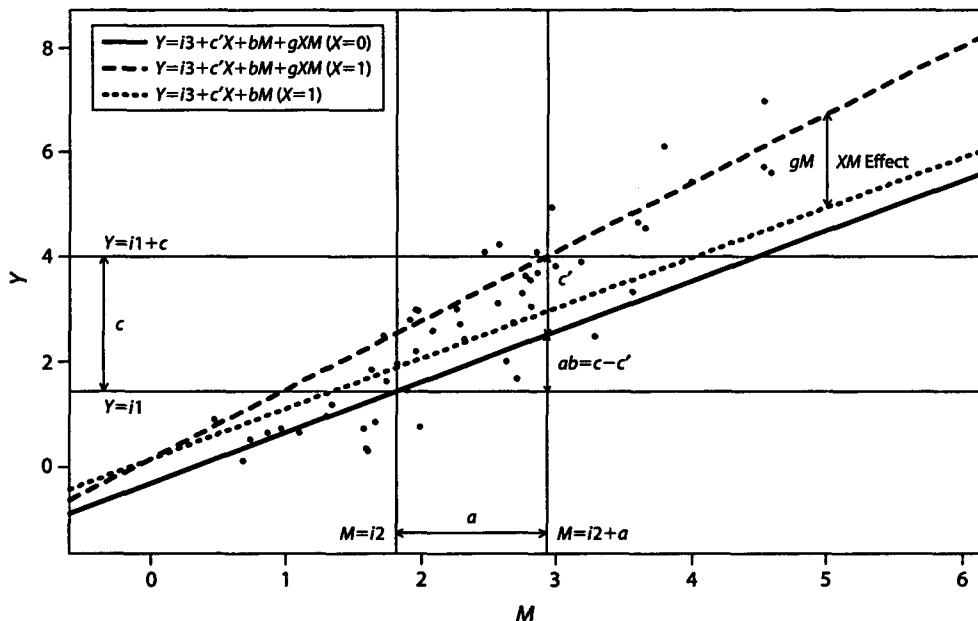

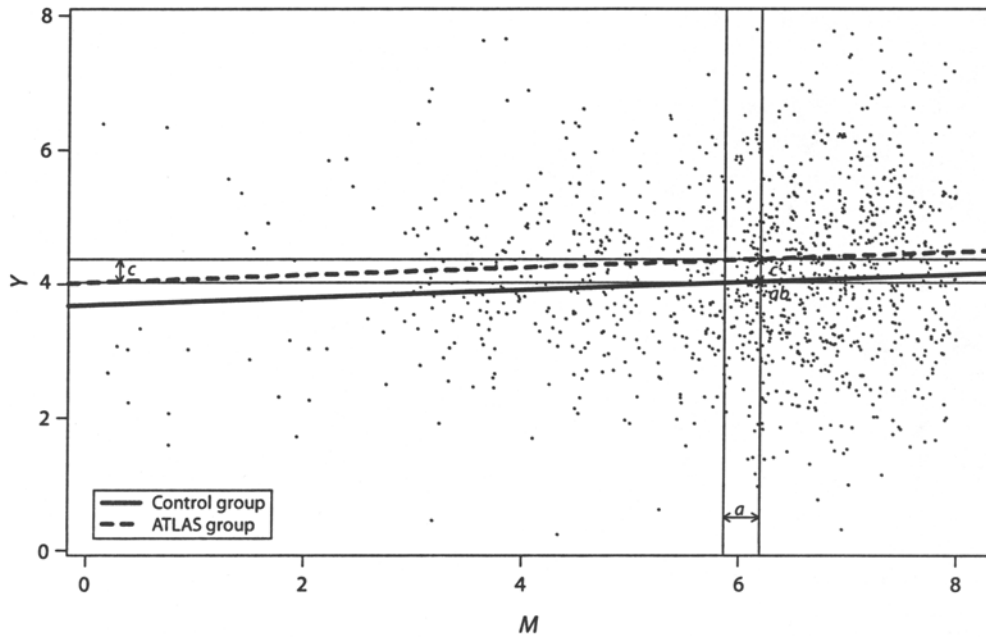**Figure 4. Plot of the mediated effect with an XM interaction.**

Figure 5. Plot of the mediated effect of exposure to the ATLAS program (X) on resistance to drug offers (M) and nutrition behaviors (Y).

diating variables, including knowledge of the effects of steroid use, peer tolerance of steroid use, and perceived severity of steroid use (see MacKinnon et al., 2001, for a complete list of the 12 mediators).

From these 12 mediator variables and 3 outcome variables, two examples are illustrated; note that the data in these plots have had a small amount of random error added to them, a process called jittering, to make them easier to read. The first example is that exposure to the ATLAS program (X) increased an athlete's ability to resist offers of steroids (M), which, in turn, increased an athlete's healthy nutrition behavior (Y). Using Cohort 1, wave C of the ATLAS data gives estimates and standard errors for the parameters of Equations 1–3 of $\hat{a} = 0.33$ ($SE = 0.0636$), $\hat{b} = 0.058$ ($SE = 0.0640$), $\hat{c} = 0.347$ ($SE = 0.0637$), and $\hat{c}' = 0.33$ ($SE = 0.0238$), which are plotted in Figure 5. The plot shows the difference in magnitude between the direct effect $\hat{c}$ and the indirect effect $\hat{a}\hat{b}$, with the direct effect making up a greater proportion of the overall effect, illustrated by the two horizontal lines. Looking at this plot, it can be seen that the proportion of the effect of X on Y mediated by M is very small and not likely to be significant. Testing for mediation using the PRODCLIN program (MacKinnon, Fritz, et al., 2007) to create a confidence interval for $\hat{a}\hat{b}$ gave a 95% confidence interval of [−0.02161, 0.06354], which shows that the indirect effect is not statistically significant.

The second example is that exposure to the ATLAS program (X) decreased an athlete's belief in media advertisements about the positive aspects of steroids (M), which then increased the athlete's strength training self-efficacy (Y). Using Cohort 1, wave C of the ATLAS data gives estimates for the parameters of Equations 1–3 of $\hat{a} = -0.68$ ($SE = 0.0745$), $\hat{b} = -0.23$ ($SE = 0.0231$), $\hat{c} = 0.31$ ($SE =$

0.0620), and $\hat{c}' = 0.16$ ($SE = 0.0619$), which are plotted in Figure 6. Looking at this plot, we see that the proportion of the total effect, $\hat{c}$, that is attributable to the mediator is much larger, around 60%, compared with the plot in Figure 5. Testing the significance of the mediated effect using the PRODCLIN program (MacKinnon, Fritz, et al., 2007) to create a confidence interval for $\hat{a}\hat{b}$ gives a 95% confidence interval of [0.11332, 0.20432], which leads to the conclusion that the mediator is significant.

Summary and Example Computer Programs

The plots discussed here are for the single mediator model only. However, many mediation models are much more complex, with multiple mediators or longitudinal data. Plotting these larger models requires plots more complex than the plots presented here, which are therefore harder to interpret. For instance, in the multiple mediator case, the mediator variables would have to be standardized to place them on the same metric, and then the plots of the individual mediators would have to be laid over one another. The interpretation of such a plot would then depend on the number of mediators and the signs and magnitudes of the individual mediated effects.

To help researchers in creating mediation plots for their own data, sample code for two common statistics packages, R (R Core Development Team, 2006) and SAS (SAS Institute, 2006), are included in the Appendix, as are instructions for loading and using the code. Electronic versions of these programs are available at www.public.asu.edu/~davidpm/ripl/mediate.htm. These sample programs are meant to act as an outline for researchers who are plotting single mediator models. Researchers are encouraged to modify the programs as necessary to explore their data and to produce plots that fit their individual needs.
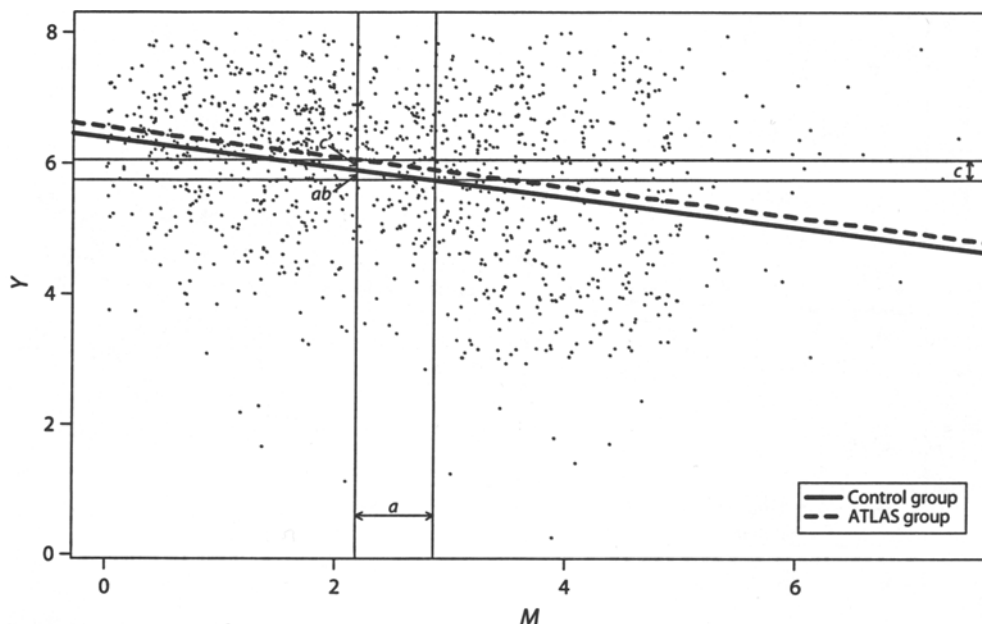
Figure 6. Plot of the mediated effect of exposure to the ATLAS program (X) on belief in media advertisements (M) and strength training self-efficacy (Y).

AUTHOR NOTE

REFERENCES

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

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

CHEN, H. T. (1990). *Theory-driven evaluations.* Thousand Oaks, CA: Sage.

FRITZ, M. S., & MACKINNON, D. P. (2006, June). *Plots of the XM interaction in models of the mediated effect.* Poster presented at the annual meeting of the Society for Prevention Research, San Antonio, TX.

GOLDBERG, L., ELLIOT, D. [L.], CLARKE, G. N., MACKINNON, D. P., MOE, E. [L.], ZOREF, L., ET AL. (1996). Effects of a multidimensional anabolic steroid prevention intervention: The Adolescents Training and Learning to Avoid Steroids (ATLAS) program. *Journal of the American Medical Association,* **276,** 1555-1562.

GOLDBERG, L., MACKINNON, D. P., ELLIOT, D. L., MOE, E. L., CLARKE, G. [N.], & CHEONG, J. (2000). The Adolescents Training and Learning to Avoid Steroids program: Preventing drug use and promoting health behaviors. *Archives of Pediatrics & Adolescent Medicine,* **154,** 332-338.

MACKINNON, D. P. (2008). *An introduction to mediation analysis.* Mahwah, NJ: Erlbaum.

MACKINNON, D. P., FAIRCHILD, A. J., & FRITZ, M. S. (2007). Mediation analysis. *Annual Review of Psychology,* **58,** 593-614.

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., GOLDBERG, L., CLARKE, G. N., ELLIOT, D. L.,

CHEONG, J., LAPIN, A., ET AL. (2001). Mediating mechanisms in a program to reduce intentions to use anabolic steroids and improve exercise self-efficacy and dietary behavior. *Prevention Science,* **2,** 15-28.

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,** 83-103.

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.

MACKINNON, D. P., TABORGA, M. P., & MORGAN-LOPEZ, A. A. (2002). Mediation designs for tobacco prevention research. *Drug & Alcohol Dependence,* **68,** S69-S83.

MERRILL, R. (1994). *Treatment effect evaluation in non-additive mediation models.* Unpublished dissertation, Arizona State University, Tempe.

R CORE DEVELOPMENT TEAM (2006). R (Version 2.4.0) [Computer Program]. Vienna, Austria: The R Foundation for Statistical Computing. Available at www.r-project.org.

SAS INSTITUTE (2006). SAS (Version 9.1) [Computer Program]. Cary, NC: SAS Institute, Inc.

ARCHIVED MATERIALS