Mediation Analysis: A Retrospective Snapshot of Practice and More Recent Directions

Lois A. Gelfand,
University of Pennsylvania

Janell L. Mensinger, and
Reading Hospital and Medical Center

Thomas Tenhave
University of Pennsylvania

Abstract

R. Baron and D. A. Kenny's (1986) paper introducing mediation analysis has been cited over 9,000 times, but concerns have been expressed about how this method is used. The authors review past and recent methodological literature and make recommendations for how to address 3 main issues: association, temporal order, and the no omitted variables assumption. The authors briefly visit the topics of reliability and the confirmatory–exploratory distinction. In addition, to provide a sense of the extent to which the earlier literature had been absorbed into practice, the authors examined a sample of 50 articles from 2002 citing R. Baron and D. A. Kenny and containing at least 1 mediation analysis via ordinary least squares regression. A substantial proportion of these articles included problematic reporting; as of 2002, there appeared to be room for improvement in conducting such mediation analyses. Future literature reviews will demonstrate the extent to which the situation has improved.

Keywords

causal modeling; mediation; path analysis

In 1986, R. Baron and Kenny Presented a method, commonly known as *mediation analysis*, for demonstrating that a data set is consistent with a model in which an intervening (or middle) variable helps explain how an independent variable influences a dependent variable. Researchers investigating mediation are interested in understanding the causal chain of events that explains how one affects another (e.g., how psychiatric treatments work, how social or work environments influence judgments or behavior). In their article, R. Baron and Kenny presented a simple, regression-based method requiring no specialized software, which has had a huge impact: To date, it has been cited over 9,000 times (ISI, 2008). However, there are theoretical and empirical reasons for concern about the application of this method of assessing mediation. The 1986 article focused on the distinction between moderation and mediation and understandably did not include extensive discussions about the complexities of path modeling and structural equation modeling (SEM), of which mediation analysis can be considered a special case. Thus, it is not surprising that both Cole and Maxwell (2003), in a literature survey of mediation analysis via SEM, and Frazier, Tix, and Barron (2004), in a review of 10 articles published in the *Journal of Counseling Psychology* in 2001, reported problems.

The purpose of the present article was to complement the existing literature by exploring several topics that have received relatively little attention. A number of important articles concerning the application of mediation analysis directed to psychological and psychiatric researchers have been published since 2002 (cf. Kraemer, Wilson, Fairburn, & Agras, 2002; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002a; Maxwell & Cole, 2007). It is not clear how long it should take for new literature about mediation analysis to influence practice, though it seems that some lag should be expected. Thus, for each topic we review some of the literature applicable to mediation analysis up to 2002 and document the presentation of mediation analyses in a sample of 50 peer-reviewed journal articles from 2002—recommendations made during the 16 years prior should arguably have been absorbed. Then, we summarize the more recent literature, which should prove useful to researchers intending to conduct mediation analyses and also to those interested in conducting future mediation analysis literature surveys.

Mediation analysis is an application of associational causal modeling (causation is modeled between subjects using measures of association), assuming no unmeasured variables. Other causal modeling approaches—for example, SEM with latent variables, or approaches in which causation is modeled within subjects, such as the potential-outcomes approach (cf. Holland, 1988)—are outside the scope of this article. In the associational tradition, X is said to cause Y if three conditions are met. First, variation in X is associated with variation in Y. Second, change in X temporally precedes change in Y (Judd & Kenny, 1981). Third, there are no unmeasured variables (i.e., omitted variables) that are correlated with X, affect Y, and are not causally intermediate (cf. Jöreskog & Sörbom, 1993). In a three-variable mediation model, independent variable X is hypothesized to cause mediator M, which, in turn, causes dependent variable Y, such that accounting for the effect of X on M and of M on Y explains, in part or in whole, the influence of X on Y. Thus, the three criteria for causality must apply to the relation of X to Y, X to M, and M to Y.

After presenting the method for article selection, we examine the assessment of association, temporal precedence, and the no omitted variables assumptions in mediation analyses. We then briefly examine two additional issues relevant to mediation analysis: reliability and the confirmatory–exploratory distinction. We make recommendations regarding how to address each of these issues and note how they were addressed in the published literature. Last, we discuss the implications of our findings.

Method

Selection of Articles

The first author examined 50 articles randomly selected from 410 articles from the first 9 months of 2002 (ISI, 2002) that cite R. Baron and Kenny (1986)[1] and include a mediation analysis, conducted through ordinary least squares (OLS) regression. The 410 articles were placed in random order (through an Excel random-number generator) and reviewed one by one until 50 English-language articles (in which at least one mediation analysis was reported that did not meet exclusion criteria) were collected. These 50 articles were obtained from examination of the first 109 articles. Thus, the sampling constitutes a negative binomial experiment, in which the proportion of such articles represented in the sampling frame is estimated by $(m – 1)/(n – 1)$, in which m is the number of desired cases in the sample and n is the size of the sample taken in order to achieve m; the variance is approximately $[m(n – m)]/[n^2(n – 1)]$ (Haldane, 1945). The proportion of articles containing mediation analyses

---

[1]This sample, taken from all citations citing R. Baron and Kenny (1986) with publication dates listed for the year 2002 (up until September 25, 2002), was originally obtained for the purpose of the first author's unpublished dissertation. We desired a sample of articles that would reflect mediation analysis practice as late as possible while at the same time prior to the influence of the wave of developments we observe as having begun in 2002 and which is still in progress. Because this sample served our needs, we used it out of convenience.

conducted through OLS regression in the 410 articles that could potentially have been examined is estimated to be .45 (95% confidence interval = .34, .55). A minority of journals outside of the field of psychology were represented, and in all cases psychological or behavioral constructs were examined.

Results

Article Selection

Table 1 shows the types of analyses presented in the 109 journal articles examined. Most articles (86, or 81%) contained at least one mediation analysis; the remaining 23 dealt with moderation only, or constituted literature reviews or methodological discussions. Of the 86 articles containing at least one mediation analysis, 50 (58%) used OLS regression and thus met our inclusion criteria. In the rest of the present article, we refer to the mediation analyses presented in these 50 articles.

Mediation Issues: Three Causal Conditions

Association—The association of X with Y, X with M, and M with Y is assessed, according to R. Baron and Kenny's (1986) method, by estimating three linear regression equations. If all three variables are standardized, the three sample regression equations are:

$$Y_{predicted} = cX \tag{1}$$

$$M_{predicted} = aX \tag{2}$$

$$Y_{predicted} = c\prime X + b\prime M \tag{3}$$

in which residuals e_1, e_2, and e_3 are the observed minus the predicted values of the dependent variables in Equation 1, Equation 2, and Equation 3, respectively, and are sampled from errors ε_1, ε_2, and ε_3 (each normally distributed with mean zero and uncorrelated with corresponding predictor variables; ε_2 and ε_3 are also uncorrelated). The logic of mediation applies equally to analyses using unstandardized coefficients, and when comparing coefficients across different groups using the same metrics, unstandardized coefficients are preferable (Asher, 1983; James, Mulaik, & Brett, 1982). However, to simplify this discussion, coefficients are, unless otherwise indicated, assumed to be standardized.

Hypothesized causal relations among variables in this mediation context can be depicted via path diagrams. Path diagram coefficients can be conceptualized at one of three levels: causal (structural) weights in a hypothetical model, values of measures of association in a population (parameters), or sample values of these measures (statistics). In this article, we conceptualize path diagram coefficients in empirical terms—as measures of association (e.g., regression coefficients)—rather than as causal weights. Furthermore, symbols representing these coefficients represent sample estimates rather than population values so that they are analogous to results from data analysis. Table 2 shows the conventions used in this article.

Figure 1A depicts correlations between X, M, and Y that are estimated to make inference on the three-variable model in Figure 1B in which X causes Y both directly and indirectly (through M).[2] In more complex models of such relations (e.g., with covaried baseline scores for M or Y), the paths between X, M, and Y are not estimated by the same simple and partial regression

coefficients depicted, but in every case there are coefficients that correspond to a, b', c, and c', so we use this notation to represent either case. Because of the linear model framework and resulting algebraic relations among regression coefficients (Wright, 1934), $c = c' + ab'$. If a model such as that depicted in Figure 1B is true, the empirical coefficients estimate causal effects as follows: c' estimates the direct effect of X on Y, not acting through M, and ab' estimates the indirect effect of X on Y through M. The total effect of X on Y is defined as the sum of the direct and indirect effects; c estimates the total effect of X on Y.[3]

R. Baron and Kenny (1986) described mediation analysis in four steps. Step 1 involves testing the significance of c to determine that there is a relation to be mediated. If significant, one tests the significance of a (Step 2) to demonstrate a relation between X and M. In Step 3, a significant b' shows that there is a relation between M and Y not accounted for by X. Once Steps 2 and 3 are passed, evidence consistent with a nonzero indirect effect has been obtained; the model is consistent with either partial mediation, complete mediation, or suppression.[4] In the first part of Step 4, which we call Step 4a, the observed values of c and c' are compared; if c' is smaller than c, the data are consistent with mediation; if c' is larger than c, the data are consistent with suppression. No significance test is necessary for this step. If Step 4a is passed, one examines the significance of c' to determine if the data are consistent with partial versus complete mediation (in what we call Step 4b); if c' is smaller than c but significantly different from 0, the data are consistent with partial mediation. If c' is smaller than c but not significantly different from 0, the data are consistent with complete mediation. These steps are summarized in Table 3.

Concern 1: Hypothesis Testing

Steps 2 and 3 test the associations a and b', whose values are both nonzero if the mediation model is true. Because these associations make up the purported indirect effect, the significance of a and b' jointly is sufficient statistical evidence for the plausibility of intervening variable effects. However, ambiguities in R. Baron and Kenny's (1986) report could lead readers to believe that an additional test is necessary to determine whether or not mediation is plausible. We examine three problematic methods for conducting a mediation analysis that uses an additional test.

Significant–Nonsignificant method—One method seemingly supported by R. Baron and Kenny (1986) is to test the difference between c and c' by comparing the significance of c' with that of c and conclude that mediation is plausible when c is significant and c' is not. There are several problems with this method. First, a reduction in significance from c to c' fails to demonstrate that a difference between c and c' is significant. Second, by ruling out mediation whenever c' is significant, this method fails to allow for a model of partial mediation. Last, because of collinearity between M and X, it is possible for c' to be equal to or larger than c (indicating no mediation or possible suppression, respectively) when c is significant and c' is not. This scenario may lead to an incorrect conclusion that mediation is present.

---

[2]Although R. Baron and Kenny (1986) did not distinguish complete mediation from partial mediation diagrammatically, Figure 1B does not, according to SEM convention, depict complete mediation; the path drawn from X and Y indicates that the path coefficient purportedly estimated by c' is nonzero (James et al., 1982). In addition, as MacKinnon, Krull, & Lockwood (2000) noted, this diagram does not necessarily depict partial mediation, as the signs of the paths are not indicated; patterns of nonzero paths consistent with suppression could also be represented.

[3]If the hypothesized model is not true, the corresponding effects may not exist, even though the coefficients of interest can be calculated from observed data and subjected to significance testing.

[4]No instructions are given regarding how to proceed if suppression is hypothesized or if evidence for suppression is unexpectedly found. Although not consistent with mediation as defined by R. Baron and Kenny (1986), it may be interesting in its own right. Furthermore, we note that, in certain cases, effects in opposite directions can cancel one another out, such that the correlation between X and Y is essentially zero, and Step 1 of the R. Baron and Kenny (1986) method would not be passed. See Davis (1985), MacKinnon et al. (2000), and Shrout and Bolger (2002) for discussions of suppression or inconsistent mediation.

Redundancy method—A second method, inferable from the inclusion of a variant of the Sobel (1982) test in R. Baron and Kenny's (1986) article, as well as statements made by Sobel (1982, 1987), involves asserting that (a) it is necessary to perform a test of ab' once R. Baron and Kenny's method has been completed and (b) this test must be significant for mediation to be supported. However, we know of no evidence demonstrating the logical superiority of an ab' test over the joint test of a and b' in demonstrating the plausibility of a mediation model; the null hypotheses of both approaches are consistent with the presence of a purported indirect effect. Furthermore, the Sobel (1982, 1987) methods assume that a sufficiently large sample is used to invoke the assumption of a normal sampling distribution for the product ab', which tends to be skewed for small samples. Bollen and Stine (1990) found that, in at least one case, a sample of 173 is sufficient, but a sample of 50 is not, and the findings of Stone and Sobel (1990) suggest that a sample of at least 200 is preferable. However, Sobel (1982, 1987) did recommend constructing a confidence interval for the purported indirect effect. This is something for which a single-parameter test is required; the ability to provide a confidence interval represents a strong advantage to such tests over the joint test of a and b'.

Partial correlation method—A third method that had appeared in the literature (cf. J. Baron, Hershey, & Kunreuther, 2000) involves comparing the partial correlation of X and Y (controlling for M, which we designate c'') to the zero-order correlation of X and Y (i.e., comparing c'' to c; cf. Olkin & Finn, 1990; Steiger, 1980) to demonstrate support for a mediated effect. Although the partial correlation coefficient c'' can be close in value to the standardized partial regression coefficient c' (e.g., when $a = b$, $c'' = c'$),[5] it can also be considerably different. Tests of the difference between c and c'' can lead to substantially different conclusions than can tests of the difference between c and c'. For example, c'' can be equal to c even when c' is considerably smaller than c, so that a valid test of the purported indirect effect would demonstrate consistency with mediation, whereas the partial correlation test would not. The reverse can also occur; the partial correlation test can suggest that data are consistent with mediation, whereas a valid test would show that they are not.

Findings from the sample—Table 4 shows examples of phrases used to determine that one of the three problematic methods was used. It was possible for more than one problematic method to be used within the same analysis. For instance, the example illustrating the partial correlation method also demonstrated the significant–nonsignificant method. Of the 50 articles, more than half (54%) included one or more of the three problematic methods for testing the indirect effect. The majority of such articles (17) included the significant–nonsignificant method, followed by articles (11) containing the redundancy method. Use of the partial correlation method accounted for a relatively small number (4) of such articles. There were 40 analyses from 14 articles in which both R. Baron and Kenny's (1986) method, with its joint test of a and b', and an ab' test were used; for the ab' test, authors cited either R. Baron and Kenny or Sobel (1982,1988). In most cases, the joint a and b' and ab' test results agreed. Where inconsistent, the joint test indicated either a significant mediation effect or a significant a with b' at trend level, whereas the ab' test was at trend level or nonsignificant, respectively. With one exception ($n = 1,090$), sample sizes for these analyses ranged from 32 to 150; in eight studies, the sample size was 50 or fewer. No confidence intervals for purported indirect effects were reported.

Directions since 2002—Holmbeck (2002) criticized what we refer to as the significant–nonsignificant method, although at the same time promoting what we refer to as the redundancy method. Although we do not favor the reflexive addition of the Sobel (1988) test, as recommended by Holmbeck, or any other ab' test, we support the use of additional tests on the

---

[5]$c'' = (c - ab)/\{\text{sqrt}[(1 - a^2)(1 - b^2)]\}$. $c' = (c - ab)/(1 - a^2)$. If $a = b$, $c'' = (c - ab)/\{\text{sqrt}[(1 - a^2)(1 - a^2)]\} = (c - ab)/(1 - a^2) = c'$.

basis of the utility for constructing a confidence interval, or on power and Type-I error-rate considerations. In a simulation study of three classes of indirect effect tests, MacKinnon et al. (2002a) found that no one test or class of tests examined is superior to all others under all conditions. In particular, variants of the Sobel (1982) test turn out to be less powerful than the joint significance test.[6] Thus, although additional tests may be useful, once the joint significance of a and b' has been assessed it does not make sense to treat an ab' test as more definitive across the board. The weaker results observed in our sample of the ab' tests could be because of either the greater power of the joint a and b' test or the excessive Type-I error of this test under certain conditions (cf. MacKinnon et al., 2002a). Shrout and Bolger (2002) recommended a bootstrap approach for sample sizes less than 200, and MacKinnon, Lockwood, & Williams (2004) compared the performance of three single-sample methods and six resampling methods. Preacher and Hayes (2004) provided Statistical Package for the Social Sciences (SPSS) and Statistical Analysis System (SAS) macros for applying one of the resampling methods, and a SAS program to apply five of these methods is available on request from Jason Williams (RTI International, jawilliams@rti.org). We suggest that researchers consult MacKinnon and colleagues (2002a, 2004) to determine which tests should have the best statistical properties for their data and use one of the available programs if a resampling method is desired. Finally, we recommend that researchers use the partial correlation method with caution. Although under some circumstances this method has desirable statistical properties, its null hypothesis is "distinctly different" (MacKinnon et al., 2002a, p. 88) from that of other intervening variable tests. That is, the partial correlation test does not directly assess the significance of purported direct or indirect effects.

Concern 2: Examining the Sizes of Regression Coefficients

The values of a, b', c and c' in each mediation analysis presented should be reported along with their significance for technical and conceptual reasons. In an analysis with no missing data and with the same set of covariates controlled for in all three equations, $c = c' + ab'$. A technical concern arises from this equivalence; it is simple to check it using the coefficient values, but impossible unless the coefficients are reported in the manuscript. Rounding can cause small discrepancies,[7] but substantial differences between c and $c' + ab'$ are likely to represent other influences, such as inconsistently applied covariates or missing data. Covariates should be applied consistently, and the discrepancy between c and $c' + ab'$ should be examined before publication. Substantial discrepancies caused by missing values could be resolved by using the same observations to calculate a, b', c, and c', as occurs with the use of covariance structural modeling software. This could be accomplished by including only observations with no missing data or by imputing values for missing data. However, one should be aware of additional potentially unfeasible assumptions when using these procedures (Little & Rubin, 1987). We call the difference between c and $c' + ab'$ the c discrepancy and show its distribution.

The conceptual issues center on the descriptive information provided by the coefficients. Coefficient values can be used to judge the consistency of the data with hypothesized relations (Bollen, 1990; Hayduk, 1987), evaluate the practical importance of results, or simply describe

---

[6]Preacher and Hayes (2004) incorrectly stated that MacKinnon et al. (2002a) found variants of the Sobel (1982) test to be " superior in terms of power" (p. 718). The Sobel (1982) test and its variants are grouped with the joint significance of a and b' in terms of power and Type-I error. In that group, "[t]he joint significance test … appears to be the best test … as it has the most power and the most accurate Type I error rates in all cases compared to the other methods" (MacKinnon et al., 2002a, p. 99); performing Step 1 of the R. Baron and Kenny (1986) method and then substituting the Sobel (1982) test (or a variant thereof) for Steps 2 and 3, as Preacher and Hayes (2004) suggested, would result in a decrease rather than an increase in power.

[7]Rounding alone should not account for discrepancies reported in this article (see Figure 2A) larger in absolute value than .02. First, one can assume that coefficients are standardized and rounded to the nearest .01 (the one unstandardized case represented had a discrepancy of −0.03, and all other coefficients represented were smaller in absolute value than 1.00). To maximize rounding error, add .005 to a, b', and c' and calculate c: $c' + .005 + (a + .005)(b' + .005) = c' + ab' + .005a + .005b' + .005025$. To maximize .005$a$ and .005b', set a and b' equal to their maximum value 1, to get $c' + ab' + .015025$. Then, round c down to $c − .005$ to get a maximum rounding discrepancy of .020025.

the strength of observed relations. This information cannot be conveyed purely by the results of significance tests. Even if data are not found to be consistent with mediation because of nonsignificant tests, examining the coefficients a and b' allows one to assess whether the relation of, for example, a treatment to a proposed mediator was weaker than hypothesized, the relation between the proposed mediator and outcome was weaker than hypothesized, or both (cf. MacKinnon, Taborga, & Morgan-Lopez, 2002b). When data support mediation through significant mediation tests, the magnitude of the indirect effect is important, as reflected in the estimates of the product of a and b or $c - c'$.

Ideally, the expected strengths of relations should be discussed before analysis takes place (Boomsma, 2000), compared with obtained coefficient values, and used to interpret results. The magnitude of c' should be used to qualify conclusions drawn about the consistency of data with partial or complete mediation. For example, a large but nonsignificant c' may suggest that partial as well as complete mediation is plausible, and a tiny but significant c' may suggest that data are consistent with complete mediation for all practical purposes. Confidence intervals can also be useful for these judgments. Unstandardized coefficients and corresponding confidence intervals, or the information to derive them, should be reported so that results can be assessed in terms of clinical or practical significance and compared across samples.

The proportion mediated (Alwin & Hauser, 1975; Sobel, 1982) can be used to summarize the results of a mediation analysis with a single value ($ab'/[c' + ab']$). MacKinnon, Warsi, & Dwyer (1995) showed the proportion mediated to be unstable; thus, we do not recommend giving it much weight when drawing conclusions. However, we use it to succinctly summarize the results for each analysis in which coefficients are reported.

Findings from the sample—Of the 50 articles, 26 (52%) did not report all the relevant coefficients (a, b', c, c') for any analysis. The 24 articles reporting all coefficients for at least 1 analysis provided 51 such analyses. For these 51 analyses, Figure 2A is a stem-and-leaf plot showing the distribution of the c discrepancy, $c - (c' + ab')$. The stem represents the 10th place of each c discrepancy, and the leaves represent the 100th place. Thirty of the c discrepancies are 0, but some are considerably larger, suggesting influences other than rounding.

The proportion mediated could be calculated for seven additional analyses (it does not require c); Figure 2B shows the distribution of the proportion mediated ratio $ab'/(c' + ab')$ for the total of 58 analyses, using a stem-and-leaf plot with the same units as in Figure 2A. As seen in Figure 2B, lower proportions mediated tended to be associated with nonsignificant indirect effects, but there was some overlap. One analysis with a proportion mediated of 0.05 was reported to be consistent with partial mediation, whereas one analysis with a proportion mediated of 1.01 was reported as consistent with (complete) mediation only at the trend level. Although it is impossible to make generalizations about the size of a proportion mediated that should be considered important across specific contexts, it is clear from Figure 2B that one cannot judge the importance of a putative mediated effect with significance tests alone; coefficient values must be examined as well.

In all, 4 articles included unstandardized regression coefficients, 20 contained clearly identified standardized coefficients, and 21 others contained coefficients that might have been standardized regression coefficients (values were all smaller than one), but that were not clearly identified by the authors. Three articles contained only partial correlation or other coefficients, and two articles contained no coefficient values.

Directions since 2002—Cumming (2008) demonstrated that, under typical circumstances, confidence intervals provide substantially more information than do p values (although the two are statistically tied to each other) and concluded that the confidence interval is much more

useful than is the determination of whether or not a result is statistically significant. These findings underscore our recommendation to focus on coefficient values and further highlight the importance of reporting confidence intervals. Cumming also recommended withholding judgement about results before replications, meta-analyses, and other forms of converging evidence have accumulated.

Temporal Order

Implicit in the specification of a causal model is a particular temporal order of the variables (cf. Judd & Kenny, 1981). Effects cannot temporally precede causes; if X causes M and M, in turn, causes Y, then X must temporally precede M, which, in turn, must precede Y. Because an event that occurs before a second event does not necessarily cause it, the temporal intermediacy of M is a necessary but not sufficient condition for causal intermediacy. In practice, it is not always possible to know the order in which measured events occurred. Thus, a temporal order and, correspondingly, a causal order other than the one hypothesized may be plausible. Even if a model depicting the preferred causal order produces more desirable coefficient values than does a model with an alternative causal order, the model with the alternative order may more accurately describe reality.

In the simplest case, X, M, and Y each have a clear "start" time, "before which no score can exist," and "freeze" time, "after which no change in score is possible" (Davis, 1985, p. 11). Additionally, the start time of M occurs no earlier than does the freeze time of X, and the start time of Y occurs no earlier than does the freeze time of M. In fields of psychology, relations of interest are unlikely to resemble the simplest case. Although a researcher may posit that X causes M, which, in turn, causes Y, it may be impossible to demonstrate that variables are not changing simultaneously, influencing one another reciprocally, or causing one another in reverse of what is hypothesized. Gollob and Reichardt (1987, 1991) pointed out that the strength of an observed indirect effect among continuously changing variables depends on when the intermediate variable M is measured in relation to when X and Y are measured; they recommended that typical estimates of mediation effects, in which M is considered at a particular time point, be described as "time-specific" (Gollob & Reichardt, 1991, p. 253) and recognized as lower bounds for the sizes of overall effects. Last, when variables change continuously over time, a failure to take into account auto-regressive effects (the prediction of the value of a variable at one time by the value of that variable at an earlier time) can seriously bias coefficients (Gollob & Reichardt, 1987, 1991).

Different study designs offer differing levels of protection against temporal indeterminacy. Cross-sectional data offer the least information about temporal order; all evidence must come from outside the data and alternative models cannot generally be ruled out. In addition, coefficients reflect a mixture of different time-specific and auto-regressive effects, making them difficult to interpret.

Studies in which measures are made on at least two different occasions have some potential for allowing one to assess how events unfold over time, though this information may be limited. In particular, single time-point measures of continuously changing variables can make the direction of change over time difficult to assess and can also lead to biased coefficients because of the failure to model auto-regressive effects.

For experiments in which X is the experimental condition, the assignment to a level of X is a change that can be located in time, before relevant changes in M and Y. Also, because assignment to X does not change over time, there are no auto-regressive effects of X to consider. Another advantage to experiments is that, if frequent assessments are incorporated, they can allow the assessment of meaningful and replicable time-specific effects (e.g., in a randomized clinical trial M can be assessed at, say, the third treatment session). However, the temporal

precedence of M and Y must still be addressed, as well as respective auto-regressive effects. Including earlier scores of variables as covariates is recommended for reducing the bias from auto-regressive effects (Reichardt & Gollob, 1986). Ideally, the temporal order between M and Y should be based on the changes in these variables using nonoverlapping periods of time to reduce temporal ambiguity (cf. DeRubeis & Feeley, 1990; Tenhave et al., 2007).

Researchers should discuss the hypothesized temporal relations of constructs and the steps that have been taken to maximize the possibility of capturing the correct temporal effects, or, if such steps have not been taken, they should consider whether competing temporal models are plausible. Ideally, researchers should estimate coefficients for plausible alternative models. Researchers presenting experimental data should consider the temporal order of M and Y; researchers presenting time-lagged observational data should discuss the temporal order of all three variables; and researchers presenting cross-sectional data should acknowledge that alternative temporal orders have not been ruled out. In addition, autoregressive effects, when hypothesized, should be included in the model.

Findings from the sample—Twelve (24%) papers contained mediation analyses involving randomized experimental conditions. Although the temporal order of M and Y is generally undetermined, in only two cases did authors note that models with alternative temporal orders might exist. There were 11 (21%) articles analyzing observational time-lagged data. In only three cases did authors note that models with alternative temporal orders might exist. Twenty-seven (54%) articles involved the cross-sectional collection of X, M, and Y; in 12 cases, authors noted that reverse temporality among variables was possible. In total, in only 17 articles did authors note the possibility of alternative temporal orders.

Directions since 2002—Kraemer et al. (2002) and Kazdin (2007) recommended that randomized clinical trials routinely include analyses to identify mediators.[8] Stice, Presnell, Gau, & Shaw (2007) advocated making the temporal precedence of M in relation to Y a more explicit criterion of mediation analysis; Kazdin also underlined the importance of "establishing a timeline" (p. 8). Furthermore, recent work by Maxwell and Cole (2007) suggested that mediation analysis coefficients for data obtained cross-sectionally from continuously changing variables are hopelessly biased by unmodeled auto-regressive effects; mediation analyses using such data should be treated with the highest degree of skepticism.

No Omitted Variables Assumption

It is well known that two variables that are correlated are not necessarily causally related (correlation does not imply causation). A third variable may cause both the so-called independent and dependent variables (causing a spurious relation) or may be correlated with the independent variable and affect the dependent variable without being causally intermediate (a confound). The omission of such third variables from mediation analyses can bias model-based estimates (Jöreskog & Sörbom, 1993; Judd & Kenny, 1981). Thus, even if the temporal order of the three variables of interest can be convincingly established, researchers should consider the possibility that other variables have been omitted from the model that would substantially change the results.

We have already noted that assignment to levels of X places X earlier in time than do outcomes M and Y. In addition, with random assignment, characteristics influencing outcome do not systematically differ between conditions, and third variables occurring prior to or concurrently with X cannot cause both assignment to X and to levels of outcomes. Thus, in randomized

---

[8]Kraemer et al. (2002) also proposed changing the definition of mediation so that it can include statistical interaction in the absence of a treatment effect. A discussion of the merits of this proposal is outside the scope of this article.

experiments, the no omitted variables assumption is well justified as applied to the relations between X and M and between X and Y. If an association is found between X and M or X and Y (i.e., if treatment differences are observed), all three conditions for causality are met for these two cases. Thus, random assignment to experimental conditions allows the strongest possible causal inference to be made regarding X as a cause of M and Y.

However, even in a randomized experiment, subjects are not randomly assigned to M; even if M occurs before and is associated with Y, the relation between M and Y may be confounded or spurious, and mediation may not be present. In observational studies, relations between X and M and between X and Y may be confounded or spurious as well. Furthermore, when variables are examined that cannot be manipulated (e.g., sex, race) and thus not described within a potential-outcomes framework, the term causal modeling may not be justifiable (Rubin, 1986; Wilkinson, 1999).

The no omitted variables assumption cannot be verified using any data-analysis procedure. As this assumption is fundamental to causal modeling, it should be acknowledged and some statement regarding its plausibility made. If there are obvious potentially important omitted variables, their potential effects on coefficients should be assessed (James, 1980) and, if possible, included in subsequent analyses as predictor variables. Generally, researchers should be as careful when inferring cause from correlation in the context of mediation analysis as they would be in other contexts. The possibility of follow-up studies involving experimental manipulation of M, or other designs that would allow one to make stronger causal inferences, should also be considered.

Findings from the sample—In only seven articles (14%) did authors mention the possibility that omitted variables or spurious relations among modeled variables affected results.

Directions since 2002—Although the no omitted variables assumption cannot be verified, recent developments suggest that in some mediation analysis contexts it may be evaluated to a certain extent using the potential-outcomes framework and additional baseline covariate information (cf. Mealli & Rubin, 2003; Tenhave et al., 2007). Alternatively, Kraemer and colleagues (Kraemer, Kieman, Essex, & Kupfer, 2008; Kraemer et al., 2002) proposed changing the definition of mediator and mediation, such that the no omitted variables assumption and the attending causal connotations are dropped and *mechanism* is reserved to describe a causally intervening variable.

Other Issues

Reliability—One of the assumptions underlying OLS regression analysis is that predictor variables are measured without error; unreliable measurement in predictors can bias regression coefficients (Hildebrand, 1986). Estimates of the reliability of measures should be reported. In experiments, subjects are randomly assigned to levels of X; measurement of X is consequently reliable, and *c* is an unbiased estimate of the relation between X and (unstandardized) Y. (Similarly, *b* is an unbiased estimate of the relation between X and unstandardized M.) Measurement unreliability in M would be expected to lead to underestimating *b′* and thus *ab′*, and overestimating *c′* (R. Baron & Kenny, 1986; Hoyle & Kenny, 1999). In observational studies, although measurement unreliability in the independent variable and mediator would be expected to reduce *c* and *a*, respectively, effects on *b′* and *c′* are more complex and difficult to predict (R. Baron & Kenny; Cohen & Cohen, 1983). Standardized regression coefficients can also be biased by unreliable measurement in Y. Ideally, measurements should be reliable. When measurements are not reliable, the potential bias in coefficients should be noted.

Findings from the sample—Where reliability was reported, many measures were quite reliable, with Cronbach's alpha, test–retest reliability, or interrater reliability greater than .80. The lowest reliability reported was between .50 and .59 in 6 cases, between .60 and .69 in 13 cases, and between .70 and .79 in 9 cases. It was .80 or greater in 10 cases. Twelve articles did not contain information about measurement reliability. The authors of only one article noted the possibility that measurement unreliability may have biased coefficient values.

Directions since 2002—Hoyle and Robinson (2004) demonstrated that even small deviations from perfect reliability in a putative mediator can substantially bias results and recommended either restricting analyses to measures with documented reliabilities of at least .90 or using latent variable models. Trafimow (2006) described a concern for the construct validity of measures that is roughly analogous to that raised by measurement unreliability but for which there is currently no means of correction. As for unreliability, only in the case of two imperfectly valid variables is the direction of bias known, and it produces conservative estimates; he recommended avoiding correlational analyses involving more than two variables. Following this recommendation would obviously prevent one from conducting mediation analyses. However, one could accept his premises and argue that mediation analyses involving experimental manipulation of X (if X does not represent a latent construct and is thus perfectly valid), single time-point measures of M and Y, and unstandardized coefficients would predictably underestimate ab' and should be allowed.

Confirmatory–Exploratory Distinction

When applying mediation analysis, researchers should consider the degree to which their approach is confirmatory versus exploratory. This distinction refers to the decision of how many models to examine; this, in turn, involves selecting variables, arranging them in a model according to a hypothesized temporal and causal order, and choosing the strengths of the paths linking them (usually zero or nonzero). Although some researchers argue that causal analyses, including mediation analyses, should be applied in a strictly confirmatory manner (James & Brett, 1984; James, Mulaik, & Brett, 1982), others such as Boomsma (2000) allowed for more exploratory approaches, so long as the complete set of models is presented and theoretically justified in advance. At a minimum, every mediation analysis model evaluated, regardless of the number of steps passed, should be listed. Furthermore, the implications of the degree of exploration should influence the conclusions drawn (e.g., exploratory analyses require replication).

Findings from the sample—For five of the articles, it was not possible to discern the number of mediation analyses performed.

Directions since 2002—Kraemer et al. (2002) recommended that mediation analysis in the context of randomized clinical trials be treated as "hypothesis-generating" rather than "hypothesis-testing" (p. 882). Conversely, James, Mulaik, and Brett (2006) reemphasized the confirmatory nature of mediation analysis, particularly as it relates to choosing the strengths of the paths linking variables. That is, he criticized the R. Baron and Kenny (1986) method for failing to distinguish a priori between complete and partial mediation. Making this distinction would require researchers to specify in advance which model is hypothesized (using a path diagram without a linkage from X to Y to depict complete mediation), use new data to test an alternative model if the first model is disconfirmed, and use ab rather than ab' to estimate the purported indirect effect if complete mediation is supported. To the extent that not only the mediation model but also the expected magnitudes of the coefficients are specified in advance, it is possible to estimate the power to detect mediation using tables presented by Fritz and MacKinnon (2007). Trafimow (2003), using simulations based on Bayes's theorem, distinguished between research that is "(a) not theoretical or exploratory versus research that

is (b) not theoretical but is exploratory versus research that is (c) theoretical" (p. 534) by the relative degree of focus on (a) clearly rejecting a null hypothesis, versus (b) learning more about the probability that a null hypothesis is true, versus (c) learning more about the probability that a theory is true. He suggested that researchers, using the prior information available, consider the combination of prior and posterior probabilities (of findings, hypotheses, and theories) that would best serve the researchers' purposes based on this distinction, and select hypotheses to test accordingly.

Discussion

In this article, we discussed association, temporal precedence, the no omitted variables assumption, measurement reliability, and the confirmatory–exploratory distinction in mediation analysis; summarized the pre-2002 literature; showed how these issues were addressed in a 2002 literature sample; and described more recent developments. The association issues are primarily technical and, as a result, are simplest to address. Researchers who want to construct confidence intervals for or conduct additional testing of the purported indirect effect can use formulas presented in MacKinnon et al. (2002a) or access programs in Preacher and Hayes (2004) or from Jason Williams (RTI International, jawilliams@rti.org). All researchers can provide sufficient information for determining both standardized and unstandardized coefficients, verify that $c = c' + ab'$ before submitting for publication, report all coefficient values, and take the values (and confidence intervals) into account when drawing conclusions (about complete vs. partial mediation, the practical significance of the results, or reasons for failure to support mediation).

Unfortunately, although the temporal order and no omitted variables assumption are fundamental to causal modeling, there are no simple ways to address them. However, failure to acknowledge and consider the reasonableness of these underlying assumptions risks drawing conclusions that go beyond what one's data support. Mediation analysis per R. Baron and Kenny (1986) is essentially a multiple-regression analysis with Y as the criterion and X and M as predictors, something our path labeling conventions are meant to underscore. In principle, using nonoverlapping time periods to assess change allows one to rule out certain directions of temporal prediction. That is, later change cannot precede, and by extension cause, earlier change. However, simultaneous change in two variables (e.g., M and Y) cannot be ruled out using this method. In practice, measurement of change using OLS regression-based methods is a subject of some disagreement (cf. Allison, 1990; Cohen & Cohen, 1983; Rogosa & Willett, 1983).

The no omitted variables assumption is fundamental to the conceptualization of SEM, path analysis, and mediation analysis as forms of causal modeling. There seem to be three general approaches one can take regarding this assumption. First, one can try to evaluate it according to the potential-outcomes framework (cf. Mealli & Rubin, 2003; Tenhave et al., 2007). Second, one can take an SEM approach and make a case regarding its plausibility on the basis of other variables included in the model (cf. James, 1980). Last, one can embrace an approach that is not predicated on a causal modeling conceptualization and use results to formulate causal hypotheses that are tested in future studies, rather than draw causal conclusions (cf. Kraemer et al., 2008). What one should not do is ignore the no omitted variables assumption and at the same time make causal claims or statements that rest on them (e.g., policy recommendations).

Reliability of measurement is generally an issue in multiple-regression analysis, as well as in mediation analysis and other forms of associational causal modeling. Again, there seem to be three general approaches to this issue. First, one can use highly reliable measures. Second, one can use less reliable measures and attempt to model the measurement structure using SEM with latent variables. Although this approach has its appeal, we note that there remains the danger

that, because modeling relations involving latent variables using currently available software is technically easy, it may be tempting to do so without paying adequate attention to the additional assumptions entailed. Last, one can use less reliable measures without modeling the measurement structure and be additionally cautious in interpreting the values of the coefficients reported.

We note that the randomized-experiment context, in which X is the assignment variable, confers several advantages regarding the temporal precedence, no omitted variables assumption, and reliability issues. The temporal precedence of X (but not M) is established, the no omitted variables assumption is justified for the relation of X to M and of X to Y (but not of M to Y), and the measurement reliability and, in some cases, the construct validity of X is perfect.

Regardless of whether one takes a more confirmatory or exploratory approach to mediation analysis, one can make the approach explicit, and one can report results for all mediation analyses attempted. Furthermore, to the extent that one is interested in evaluating a psychological theory, one should select mediational hypotheses that are likely to be true if and only if the theory is true (Trafimow, 2003).

Our results suggest that in 2002 most citations of R. Baron and Kenny (1986) were in the context of mediation analysis, and most of these analyses were performed via OLS regression as presented in that article (though using standardized variables). We found that almost a quarter of the articles contained mediation analyses involving randomized experimental conditions; however, given the difficulties associated with modeling mediation with purely observational data, a higher ratio would be desirable. Just as Cole and Maxwell (2003) found in their sample of mediation analysis using SEM, we found a lack of attention to important assumptions underlying associational causal modeling (i.e., temporal precedence and the no omitted variables assumption). Similarly, we found reason to be concerned about the performance of hypothesis tests, the reporting of coefficient values, reliability of variables, and the confirmatory–exploratory distinction. We hope that the more frequent appearance of papers about mediation analysis in the psychological and psychiatric literature (a trend we expect to continue) will lead to improvements in the conduct of these analyses; future literature surveys will discover whether this is the case.

Acknowledgements

Biographies

Lois A. Gelfand is a visiting scholar in the Department of Psychology at the University of Pennsylvania. Her current research interests include mediation analysis and depression treatment process.

Janell L. Mensinger is the Director of Research at the Reading Hospital and Medical Center where she lectures in Research Methodology and Biostatistics for Continuing and Graduate Medical Education. Her research interests include the etiology and treatment of disordered eating.

Thomas Tenhave is a professor of biostatistics in the Department of Biostatistics and Epidemiology of the University of Pennsylvania's School of Medicine. His current research interests include informative dropout, causal models, mediation and moderation analyses,

treatment nonadherence, and designs and statistical analyses to accommodate patient preferences and adaptive treatment regimes.

REFERENCES

Allison PD. Change scores as dependent variables in regression analyses. Sociological Methodology 1990;20:93–114.

Alwin DF, Hauser RM. The decomposition of effects in path analysis. American Sociological Review 1975;40(1):37–47.

Asher, HB. Causal modeling. Beverly Hills, CA: Sage; 1983.

Baron J, Hershey JC, Kunreuther H. Determinants of priority for risk reduction: The role of worry. Risk Analysis 2000;20:413–427. [PubMed: 11051067]

Baron R, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology 1986;51:1173–1182. [PubMed: 3806354]

Bollen KA. A comment on model evaluation and modification. Multivariate Behavioral Research 1990;25(2):181–185.

Bollen KA, Stine R. Direct and indirect effects: Classical and bootstrap estimates of variability. Sociological Methodology 1990;20:115–140.

Boomsma A. Reporting analyses of covariance structures. Structural Equation Modeling 2000;7:461–483.

Cohen, J.; Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Vol. 2nd ed.. Mahwah, NJ: Erlbaum; 1983.

Cole DA, Maxwell SE. Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. Journal of Abnormal Psychology 2003;112:558–577. [PubMed: 14674869]

Cumming G. Replication and *p* intervals: *p* values predict the future only vaguely, but confidence intervals do much better. Perspectives on Psychological Science 2008;3:286–300.

Davis, JA. The logic of causal order. Beverly Hills, CA: Sage; 1985.

DeRubeis RJ, Feeley M. Determinants of change in cognitive therapy for depression. Cognitive Therapy and Research 1990;14:469–482.

Frazier PA, Tix AP, Barron KE. Testing moderator and mediator effects in counseling psychology research. Journal of Counseling Psychology 2004;51(1):115–134.

Fritz MS, MacKinnon DP. Required sample size to detect the mediated effect. Psychological Science 2007;18:233–239. [PubMed: 17444920]

Gollob HF, Reichardt CS. Taking account of time lags in causal models. Child Development 1987;58:80–92. [PubMed: 3816351]

Gollob, HF.; Reichardt, CS. Interpreting and estimating indirect effects assuming time lags really matter. In: Collins, LM.; Horn, JL., editors. Best methods for the analysis of change: Recent advances, unanswered questions, future directions. Washington, DC: American Psychological Association; 1991. p. 243-259.

Haldane JBS. On a method of estimating frequencies. Biometrika 1945;33:222–225.

Hayduk, LA. Structural equation modeling with lisrel: Essentials and advances. Baltimore, MD: Johns Hopkins University Press; 1987.

Hildebrand, DK. Statistical thinking for behavioral scientists. Boston: Duxbury Press; 1986.

Holland PW. Causal inference, path analysis, and recursive structural equations models. Sociological Methodology 1988;18:449–484.

Holmbeck GN. Post-hoc probing of significant moderational and mediational effects in studies of pediatric populations. Journal of Pediatric Psychology 2002;27:87–96. [PubMed: 11726683]

Hoyle, RH.; Kenny, DA. Sample size, reliability, and tests of statistical mediation. In: Hoyle, RH., editor. Statistical strategies for small sample research. Thousand Oaks, CA: Sage; 1999. p. 195-222.

Hoyle, RH.; Robinson, JC. Mediated and moderated effects in social psychological research. In: Sansone, C.; Morf, CC.; Panter, AT., editors. Handbook of methods in social psychology. Thousand Oaks, CA: Sage: 2004. p. 213-233.

ISI Web of Knowledge. 2002. Retrieved September 25, 2002, from http://apps.isiknowledge.com

ISI Web of Knowledge. 2008. Retrieved February 10, 2008, from http://apps.isiknowledge.com

James LR. The unmeasured variables problem in path analysis. Journal of Applied Psychology 1980;65:415–421.

James LR, Brett JM. Mediators, moderators, and tests for mediation. Journal of Applied Psychology 1984;69:307–321.

James, LR.; Mulaik, SA.; Brett, JM. Causal analysis: Assumptions, models, and data. Beverly Hills, CA: Sage; 1982.

James LR, Mulaik SA, Brett JM. A tale of two methods. Organizational Research Methods 2006;9:233–244.

Jöreskog, KG.; Sörbom, D. Lisrel 8: Structural equation modeling with the simplis command language. Hillsdale, NJ: Scientific Software International; 1993.

Judd CM, Kenny DA. Process analysis: Estimating mediation in treatment evaluations. Evaluation Review 1981;5:602–619.

Kazdin AE. Mediators and mechanisms of change in psychotherapy research. Annual Review of Clinical Psychology 2007;3:1–27.

Keenan T. Negative affect predicts performance on an object permanence task. Developmental Science 2002;5:65–71.

Kraemer HC, Kieman M, Essex M, Kupfer D. How and why criteria defining moderators and mediators differ between the Baron and Kenny and MacArthur approaches. Health Psychology 2008;27:S101–S108. [PubMed: 18377151]

Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. Archives of General Psychiatry 2002;59:877–883. [PubMed: 12365874]

Little, RJA.; Rubin, DB. Statistical analysis with missing data. New York: Wiley; 1987.

MacKinnon DP, Krull JL, Lockwood CM. Equivalence of the mediation, confounding and suppression effect. Prevention Science 2000;1(4):173–181. [PubMed: 11523746]

MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychological Methods 2002a;7:83–104. [PubMed: 11928892]

MacKinnon DP, Lockwood CM, Williams J. Confidence limits for the indirect effect: Distribution of the product and resampling methods. Multivariate Behavioral Research 2004;39(1):99–128.

MacKinnon DP, Taborga MP, Morgan-Lopez AA. Mediation designs for tobacco prevention research. Drug & Alcohol Dependence 2002b;68:S69–S83. [PubMed: 12324176]

MacKinnon DP, Warsi G, Dwyer JH. A simulation study of mediated effect measures. Multivariate Behavioral Research 1995;30(1):41–62.

Maxwell SE, Cole DA. Bias in cross-sectional analyses of longitudinal mediation. Psychological Methods 2007;12(1):23–44. [PubMed: 17402810]

Mealli F, Rubin DB. Assumptions allowing the estimation of direct causal effects. Journal of Econometrics 2003;112:79–87.

Olkin I, Finn JD. Testing correlated correlations. Psychological Bulletin 1990;108:330–333.

Pellegrini AD, Long JD. A longitudinal study of bullying, dominance, and victimization during the transition from primary school through secondary school. British Journal of Developmental Psychology 2002;20:259–280.

Preacher KJ, Hayes AF. SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behavior Research Methods, Instruments, and Computers 2004;36:717–731.

Reichardt, CS.; Gollob, HF. Satisfying the constraints of causal modeling. In: Trochim, WMK., editor. Advances in quasi-experimental design and analysis. San Francisco: Jossey-Bass; 1986. p. 91-107.

Rogosa DR, Willett JB. Demonstrating the reliability of the difference score in the measurement of change. Journal of Educational Measurement 1983;20:335–343.

Rubin DB. Comment: Which ifs have causal answers. Journal of the American Statistical Association 1986;81:961–962.

Shrout PE, Bolger N. Mediation in experimental and nonexperimental studies: New procedures and recommendations. Psychological Methods 2002;7:422–445. [PubMed: 12530702]

Sobel ME. Asymptotic confidence intervals for indirect effects in structural equation models. Sociological methodology 1982;13:290–312.

Sobel ME. Direct and indirect effects in linear structural equation models. Sociological Methods and Research 1987;16(1):155–176.

Sobel, ME. Direct and indirect effects in linear structural equation models. In: Long, JS., editor. Common problems/proper solutions: Avoiding error in quantitative research. Beverly Hills, CA: Sage; 1988. p. 46-64.

Steiger JH. Tests for comparing elements of a correlation matrix. Psychological Bulletin 1980;87:245–251.

Stice E, Presnell K, Gau J, Shaw H. Testing mediators of intervention effects in randomized controlled trials: An evaluation of two eating disorder prevention programs. Journal of Consulting and Clinical Psychology 2007;75(1):20–32. [PubMed: 17295560]

Stone CA, Sobel ME. The robustness of estimates of total indirect effects in covariance structure models estimated by maximum likelihood. Psychometrika 1990;55:337–352.

Tenhave T, Joffe M, Lynch K, Brown G, Maistro S, Beck A. Causal mediation analyses with rank preserving models. Biometrics 2007;63:926–934. [PubMed: 17825022]

Trafimow D. Hypothesis testing and theory evaluation at the boundaries: Surprising insights from Bayes's theorem. Psychological Review 2003;110(3):526–535. [PubMed: 12885113]

Trafimow D. Multiplicative invalidity and its application to complex correlational models. Genetic, Social, and General Psychology Monographs 2006;132(3):215–239.

Verkuyten M, Thijs J. School satisfaction of elementary school children: The role of performance, peer relations, ethnicity, and gender. Social Indicators Research 2002;59:203–228.

Wilkinson L. Statistical methods in psychology journals: Guidelines and explanations. American Psychologist 1999;54:594–604.

Wright S. The method of path coefficients. Annals of Mathematical Statistics 1934;5:161–215.
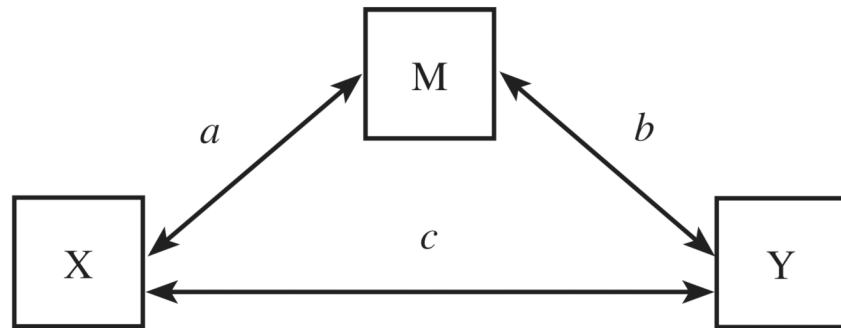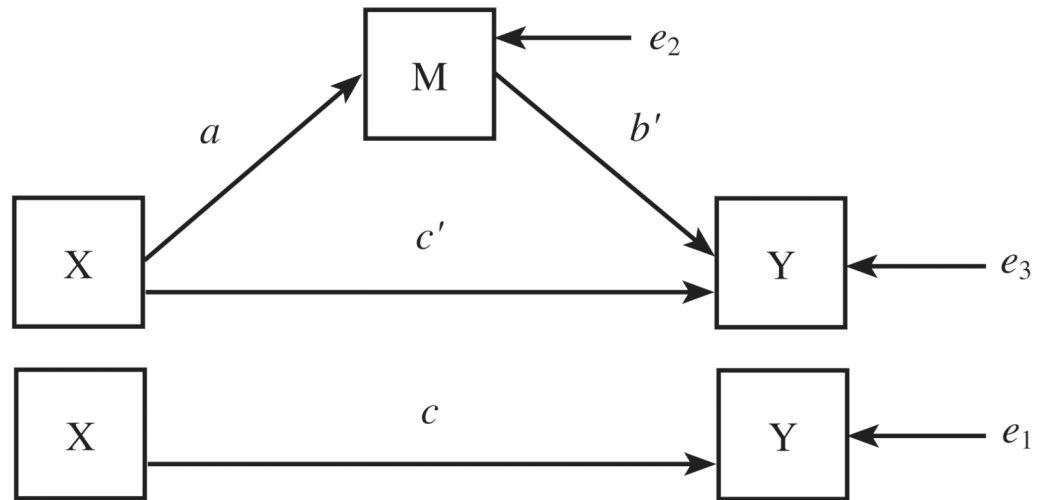
Figure 1.
Path diagrams depicting the relations among three standardized variables (X, M, and Y). Figure 1A depicts correlations. Figure 1B depicts a mediation model.

A

$n = 51$	Leaf unit = .01
Ext	(−1.03)
−0.5	62
−0.4	
−0.3	9
−0.2	
−0.1	711
(−)0	553
0.0	0000000000000000000000000000000
0.0	223456
0.1	
0.2	29
0.3	9
0.4	
0.5	2
0.6	5

B

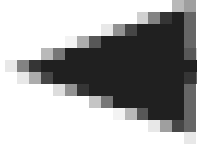|	
$n = 58$	Leaf unit = .01
Ext	(−1.43, −21.00)
0.0	56**9**
0.1	**55**99
0.2	**000**122779
0.3	0234**5**578
0.4	001**333**3457
0.5	0346
0.6	44
0.7	015
0.8	34478
0.9	1
1.0	0**1**
1.1	
1.2	
1.3	0
1.4	1
Ext	(2.10, 5.55, 15.60)

Figure 2.
Figure 2A is a stem-and-leaf plot showing the distribution of the c discrepancy [$c − (c' + ab')$] from the 51 analyses for which all four coefficients are reported. Figure 2B is a stem-and-leaf plot showing the distribution of the proportion mediated [$ab'/(c' + ab')$] from the 58 analyses in which a, b', and c' are reported. Numbers not bolded represent data that are consistent with mediation (partial or complete); bolded numbers represent data that are not consistent with mediation because either a or b' is not significant at the .05 level. Ext = extreme value.

TABLE 1
Types of Analyses Found in Sample of 109 Articles Citing R. Baron and D. A. Kenny (1986)

Method	Number of articles
≥ 1 mediation analysis	
OLS regression	50
ANCOVA	5
Logistic regression	6
Probit	1
Odds ratios	1
Non-English	1
SEM	22
No mediation analysis	
Moderation only	9
Other (e.g., literature review, methods discussion)	14

Note. OLS = ordinary least squares; ANCOVA = analysis of covariance; SEM = structural equation modeling.

TABLE 2

Path Diagram and Path Diagram Coefficient Conventions

Path coefficient symbol	Description	Path symbol
Lowercase letter (e.g., *c*)	Correlation	
Lowercase letter (e.g., *a*)	Simple standardized regression coefficient (equals correlation)	
Lowercase letter, single primed (e.g., *b′*, *c′*)	Partial standardized regression coefficient	
Lowercase letter, double primed (e.g., *c″*)	Partial correlation coefficient	Not shown
X	Independent variable	
M	Mediator	
Y	Dependent variable	
e, not enclosed in a box	Residual variable	

Note. The path diagram coefficients listed (*a*, *b′*, *c*, *c′*, and *c″*) are those that are relevant to discussing a three-variable mediation model. More complex models have corresponding coefficients. Coefficients represent sample estimates.

TABLE 3

R. Baron and D. A. Kenny's (1986) Steps

Step	Desired result
1	c is significant
2	a is significant
3	b' is significant
4a	c' is smaller than c
4b	If c' is significant, data are consistent with partial mediation; if c' is nonsignificant, data are consistent with complete mediation.

Note. If Steps 1–3 are passed, data are consistent with partial mediation, complete mediation, or suppression.

TABLE 4

Examples of Phrases Used to Determine Indirect Effect Method

Method	Phrase
Significant–Nonsignificant	"When Paths A and B were controlled statistically, Path C was no longer significant ($R^2 = .01$, $\beta = .14$); thus, the mediation hypothesis was supported, following the stipulations of [R.] Baron and Kenny (1986)." (A. D. Pellegrini & J. D. Long, 2002, p. 270)
Partial correlation	"When controlling for the effects of negative affect, the association between attention and search was reduced to non-significance ([partial correlation](33) = 0.20, $p = 0.25$)." (T. Keenan, 2002, p. 68)
Redundancy	"To begin with, school satisfaction was regressed on peer victimization, showing there is a significant and negative correlation, Beta = -0.21, $t = 7.07$, $p < 0.001$. Next, peer victimization turned out to predict social self-esteem, Beta = -0.42, $t = 15.36$, $p < 0.001$. Finally, school satisfaction was simultaneously regressed on peer victimization and social self-esteem. It turned out that social self-esteem was a significant and reliable predictor, Beta = 0.27, $t = 8.62$, $p < 0.001$, whereas the effect of peer victimization was substantially… reduced, Beta = -0.09, $t = 3.0$, $p < 0.01$. The difference between the two associations for peer victimization and school satisfaction was significant (z-value = 2.81, $p < 0.01$)." (M. Verkuyten & J. Thijs, 2002, p. 219)