

# Equivalence of the Mediation, Confounding and Suppression Effect

David P. MacKinnon,<sup>1,2</sup> Jennifer L. Krull,<sup>1</sup> and Chondra M. Lockwood<sup>1</sup>

This paper describes the statistical similarities among mediation, confounding, and suppression. Each is quantified by measuring the change in the relationship between an independent and a dependent variable after adding a third variable to the analysis. Mediation and confounding are identical statistically and can be distinguished only on conceptual grounds. Methods to determine the confidence intervals for confounding and suppression effects are proposed based on methods developed for mediated effects. Although the statistical estimation of effects and standard errors is the same, there are important conceptual differences among the three types of effects.

**KEY WORDS:** mediation; confounding; suppression; confidence intervals.

Once a relationship between two variables has been established, it is common for researchers to consider the role of a third variable in this relationship (Lazarsfeld, 1955). This paper will examine three types of third variable effects—mediation, confounding, and suppression—in which an additional variable may clarify the nature of the relationship between an independent and a dependent variable. These three concepts have largely been developed within different areas of inquiry, and although the three types of effects are conceptually distinct, they share considerable statistical similarities. Some aspects of the similarity of these concepts have been mentioned in several different articles (Olkin & Finn, 1995; Robins, 1989; Spirtes, Glymour, & Scheines, 1993; Tzelgov & Henik, 1991). In this paper, we demonstrate that mediation, confounding, and suppression effects can each be considered in terms of a general third variable model, and that point and interval estimates of mediation effects can be adapted for use in confounding and suppression frameworks. The paper focuses

on a three variable system containing an independent variable ( $X$ ), a dependent variable ( $Y$ ), and a third variable that may be a mediator ( $M$ ), a confounder ( $C$ ), or a suppressor ( $S$ ).

## MEDIATION

One reason why an investigator may begin to explore third variable effects is to elucidate the causal process by which an independent variable affects a dependent variable, a mediational hypothesis (James & Brett, 1984). In examining a mediational hypothesis, the relationship between an independent variable and a dependent variable is decomposed into two causal paths, as shown in Fig. 1 (Alwin & Hauser, 1975). One of these paths links the independent variable to the dependent variable directly (the direct effect), and the other links the independent variable to the dependent variable through a mediator (the indirect effect). An indirect or mediated effect implies that the independent variable causes the mediator, which, in turn causes the dependent variable (Holland, 1988; Sobel, 1990).

Hypotheses regarding mediated or indirect effects are common in psychological research (Alwin & Hauser, 1975; Baron & Kenny, 1986; James & Brett,

<sup>1</sup>Arizona State University.

<sup>2</sup>Correspondence should be directed to the Department of Psychology, Arizona State University, Tempe, Arizona 85287-1104. E-mail: David.Mackinnon@asu.edu

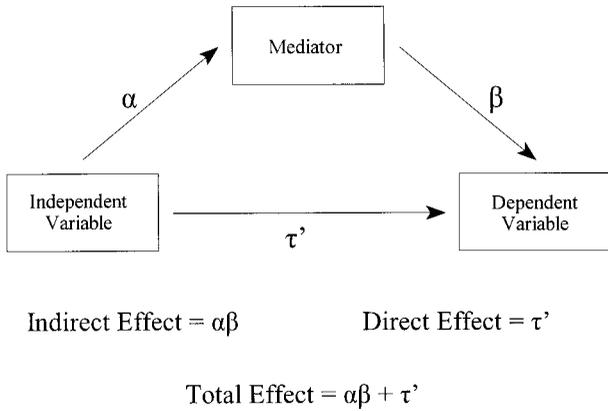


Fig. 1. A three-variable mediation model.

1984). For example, intentions are believed to mediate the relationship between attitudes and behavior (Ajzen & Fishbein, 1980). One new promising application of the mediational hypothesis involves the evaluation of randomized health interventions. The interventions are designed to change constructs, which are thought to be causally related to the dependent variable (Freedman, Graubard, & Schatzkin, 1992). Mediation analysis can help to identify the critical components of interventions (MacKinnon & Dwyer, 1993). For example, social norms have been shown to mediate prevention program effects on drug use (Hansen, 1992).

### CONFOUNDING

The concept of a confounding variable has been developed primarily in the context of the health sciences and epidemiological research.<sup>3</sup> A confounder is a variable related to two factors of interest that falsely obscures or accentuates the relationship between them (Meinert, 1986, p. 285). In the case of a single confounder, adjustment for the confounder provides an undistorted estimate of the relationship between the independent and dependent variables.

The confounding hypothesis suggests that a third variable explains the relationship between an independent and dependent variable (Breslow & Day, 1980; Meinert, 1986; Robins, 1989; Susser, 1973). Unlike the mediational hypothesis, confounding does

not necessarily imply a causal relationship among the variables. In fact, at least one definition of a confounder effect specifically requires that the third variable not be an "intermediate" variable, as mediators are termed in epidemiological literature (Last, 1988, p. 29). For example, age may confound the positive relationship between annual income and cancer incidence in the United States. Older individuals are likely to earn more money than younger individuals who have not spent as much time in the work force, and older individuals are also more likely to get cancer. Income and cancer incidence are thus related through a common confounder, age. Income does not cause age, which then causes cancer. In terms of Fig. 1, this confounding model would reverse the direction of the arrow between the independent variable and the third variable.

### SUPPRESSION

In confounding and mediational hypotheses, it is typically assumed that statistical adjustment for a third variable will reduce the magnitude of the relationship between the independent and dependent variables. In the mediational context, the relationship is reduced because the mediator explains part or all of the relationship because it is in the causal path between the independent and dependent variables. In confounding, the relationship is reduced because the third variable removes distortion due to the confounding variable. However, it is possible that the statistical removal of a mediational or confounding effect could increase the magnitude of the relationship between the independent and dependent variables. Such a change would indicate suppression.

Suppression is a concept most often discussed in the context of educational and psychological testing (Cohen & Cohen, 1983; Horst, 1941; Lord & Novick, 1968; Velicer, 1978). Conger (1974, pp. 36–37) provides the most generally accepted definition of a suppressor variable (Tzelgov & Henik, 1991): "a variable which increases the predictive validity of another variable (or set of variables) by its inclusion in a regression equation," where predictive validity is assessed by the magnitude of the regression coefficient. Thus, a situation in which the magnitude of the relationship between an independent variable and a dependent variable becomes larger when a third variable is included would indicate suppression. We focus on the Conger definition of suppression in this article, although there are more detailed discussions of sup-

<sup>3</sup>Here we are not referring to confounding discussed in the analysis of variance for incomplete designs (see Winer, Brown, & Michels, 1991 and Kirk, 1995).

pression based on the correlations among variables (see above references and also Hamilton, 1987; Sharpe & Roberts, 1997).

### COMPARING MEDIATION, CONFOUNDING, AND SUPPRESSION

Within a mediation model, a suppression effect would be present when the direct and mediated effects of an independent variable on a dependent variable have opposite signs (Cliff & Earleywine, 1994; Tzelgov & Henik, 1991). Such models are known as inconsistent mediation models (Davis, 1985), as contrasted with consistent mediation models in which the direct and mediated effects have the same sign.

The most commonly used method to test for mediation effects assumes a consistent mediation model and does not allow for suppression or inconsistent mediation. This method involves three criteria for determining mediation (Baron & Kenny, 1986; Judd & Kenny, 1981a):

1. There must be a significant relationship between the independent variable and the dependent variable,
2. There must be a significant relationship between the independent variable and the mediating variable, and
3. The mediator must be a significant predictor of the outcome variable in an equation including both the mediator and the independent variable.

McFatter (1979) presented a hypothetical situation in which an inconsistent mediation (suppression) effect is present, but which would not meet the first criterion for mediation listed above. Suppose that a researcher is interested in the interrelationships among workers' intelligence ( $X$ ), level of boredom ( $M$ ), and the number of errors made on an assembly line task ( $Y$ ). It can be plausibly argued that, all else being equal, the more intelligent workers would make fewer errors, the more intelligent workers would exhibit higher levels of boredom, and boredom would be positively associated with number of errors. Thus the direct effect of intelligence on errors would be negative, and the indirect effect of intelligence on errors mediated by boredom would be positive. Combined, these two hypothetical effects may cancel each other out, resulting in a total effect of intelligence on errors equal to zero. This nonsignificant overall relationship would fail to meet the first of the

three criteria specified above, leading to the erroneous conclusion that mediation was not present in this situation. The possibility that mediation can exist even if there is not a significant relationship between the independent and dependent variables was acknowledged by Judd and Kenny (1981b, p. 207), but it is not considered in most applications of mediation analysis using these criteria. Granted, a scenario in which the direct and indirect effects entirely cancel each other out may be rare in practice, but more realistic situations in which direct and indirect effects of fairly similar magnitudes and opposite signs result in a nonzero but nonsignificant overall relationship are certainly possible.

The notion of suppression is also present in descriptions of confounding. The classic example of a suppressor is a confounding hypothesis described by Horst (1941) involving the prediction of pilot performance from measures of mechanical and verbal ability. When the verbal ability predictor was added to the regression of pilot performance on mechanical ability, the effect of mechanical ability increased. This increase in the magnitude of the effect of mechanical ability on pilot performance occurred because verbal ability explained variability in mechanical ability; that is, the test of mechanical ability required verbal skills to read the test directions.

Breslow and Day (1980, p. 95) noted this distinction between situations in which the addition of a confounding variable to a regression equation reduces the association between an independent and a dependent variable and those suppression contexts in which the addition increases the association. They term the former "positive confounding" and the latter "negative confounding." Although defined in terms of epidemiological concepts, positive and negative confounding are analogous to consistent and inconsistent mediation effects, respectively.

### POINT ESTIMATORS OF THE THIRD VARIABLE EFFECTS

Methods of assessing third variable effects involve comparing the effect of  $X$  on  $Y$  in two models, one predicting  $Y$  from only the  $X$  variable, the other predicting  $Y$  from both  $X$  and the third variable. The third variable effect (i.e., the mediated, confounding, or suppression effect) is the difference between the two estimates of the relationship between the independent variable  $X$  and the dependent variable  $Y$ . In the discussion below, the general model is de-

scribed in terms of a third variable ( $Z$ ), which could be a mediator ( $M$ ), a confounder ( $C$ ), or a suppressor ( $S$ ). We assume multivariate normal distributions and normally distributed error terms throughout.

The effect of a third variable can be calculated in two equivalent ways (MacKinnon, Warsi, & Dwyer, 1995) based on either the difference between two regression parameters ( $\tau - \tau'$ ) mentioned above or the multiplication of two regression parameters ( $\alpha\beta$ ). In the first method, the following two regression equations are estimated:

$$Y = \beta_{01} + \tau X + \varepsilon_1, \quad (1)$$

$$Y = \beta_{02} + \tau' X + \beta Z + \varepsilon_2, \quad (2)$$

where  $Y$  is the outcome variable,  $X$  is the program or independent variable,  $Z$  is the third variable,  $\tau$  codes the overall relationship between the independent and the dependent variable,  $\tau'$  is the coefficient relating the independent to the dependent variable adjusted for the effects of the third variable,  $\varepsilon_1$  and  $\varepsilon_2$  code unexplained variability, and the intercepts are  $\beta_{01}$  and  $\beta_{02}$ .

In the first equation, the dependent variable is regressed on the independent variable. The  $\tau$  coefficient in this equation is an estimate of the total or unadjusted effect of  $X$  on  $Y$ . In the second equation, the dependent variable is regressed on the independent variable and the third variable. The  $\tau'$  coefficient in Equation 2 is an estimate of the effect of  $X$  on  $Y$  taking into account the third variable.

In a mediation context, the difference  $\tau - \tau'$  represents the indirect (i.e., mediated) effect that  $X$  has on  $Y$  by causing changes in the mediator, which then causes the dependent variable (Judd & Kenny, 1981a). In a confounding context, the difference  $\tau - \tau'$  is an estimate of confounder bias (Selvin, 1991, p. 236). The signs and magnitudes of the  $\tau$  and  $\tau'$  parameters indicate whether or not the third variable operates as a suppressor. If the two parameters share the same sign, a  $\tau'$  estimate closer to zero than the  $\tau$  estimate (a direct effect smaller than the total effect) indicates mediation or positive confounding, while a situation in which  $\tau$  is closer to zero than  $\tau'$  (a direct effect larger than the total effect) indicates suppression (or inconsistent mediation or negative confounding, depending on the conceptual context of the analysis). In some cases of suppression, the  $\tau$  and  $\tau'$  parameters may have opposite signs.

An alternative method of estimating third variable effects also involves estimation of two regression

equations. The first of these, repeated here, is identical to Equation 2:

$$Y = \beta_{02} + \tau' X + \beta Z + \varepsilon_2. \quad (2)$$

However, in this instance, attention is focused on  $\beta$ , the coefficient associated with the third variable, rather than on the estimate of the relationship between the independent and dependent variables ( $\tau'$ ). Next, a coefficient relating the independent variable to the third variable ( $\alpha$ ) is computed:

$$Z = \beta_{03} + \alpha X + \varepsilon_3, \quad (3)$$

where  $Z$  is the third variable,  $\beta_{03}$  is the intercept,  $X$  is the program or independent variable, and  $\varepsilon_3$  codes unexplained variability. The product of the two parameters  $\alpha\beta$  is the third variable effect, which is equivalent to  $\tau - \tau'$  (MacKinnon, Warsi, & Dwyer, 1995). In the confounding context, the point estimate of  $\alpha\beta = \tau - \tau'$  is the confounder effect. In those situations in which  $\alpha\beta$  has sign opposite to that of  $\tau'$ ,  $\alpha\beta$  is an estimate of the suppressor effect.

In summary, suppression, mediation, and confounding effects can be estimated by the difference between regression coefficients  $\tau - \tau'$ , which is also equal to  $\alpha\beta$ . Suppression effects can be present within either the mediational or the confounding context and are defined by the relative signs of the direct (or unadjusted) and mediated (or confounding) effects. Table 1 summarizes the possible combinations of  $\alpha\beta = \tau - \tau'$  (third variable effect) and  $\tau'$  (direct effect) population true values and lists the type of effect present in a sample. Those cases labeled "suppression" could involve either inconsistent mediation or negative confounding, depending on the conceptual framework. Those cases labeled "mediation or confounding" could actually involve either consistent mediation or positive confounding, depending on the types of variables and relationships involved. Exactly how the effect would be labeled (mediation or confounding) depends on the framework used to understand the phenomenon. Whether the effect would be termed consistent mediation/positive confounding or suppression (inconsistent mediation or negative confounding) would depend on the relationship between sample estimates of  $\tau$  and  $\tau'$ .

In Table 1, positive, negative, and zero true values for the population third variable effect, the population direct effect, and the sample values of  $\hat{\tau} > \hat{\tau}'$  and  $\hat{\tau} < \hat{\tau}'$  are considered, resulting in 18 combinations. Note that there is evidence of mediation, confounding, or suppression only in the third and sixth rows of Table 1. When the population value of the

Table 1. Interpretation of Sample Third Variable Effects Given the Sign of the Population Third Variable and Direct Effects

		Population value of the direct effect ( $\tau'$ )					
		Zero		Positive		Negative	
Population value of the third variable effect ( $\alpha\beta = \tau - \tau'$ )	Zero	$\hat{\tau} > \hat{\tau}'$	Possible by chance	$\hat{\tau} > \hat{\tau}'$	Possible by chance	$\hat{\tau} > \hat{\tau}'$	Possible by chance
		$\hat{\tau} < \hat{\tau}'$	Possible by chance	$\hat{\tau} < \hat{\tau}'$	Possible by chance	$\hat{\tau} < \hat{\tau}'$	Possible by chance
	Positive	$\hat{\tau} > \hat{\tau}'$	Complete mediation or complete confounding	$\hat{\tau} > \hat{\tau}'$	Mediation or confounding	$\hat{\tau} > \hat{\tau}'$	Suppression
		$\hat{\tau} < \hat{\tau}'$	Possible by chance	$\hat{\tau} < \hat{\tau}'$	Possible by chance	$\hat{\tau} < \hat{\tau}'$	Possible by chance
	Negative	$\hat{\tau} > \hat{\tau}'$	Possible by chance	$\hat{\tau} > \hat{\tau}'$	Possible by chance	$\hat{\tau} > \hat{\tau}'$	Possible by chance
		$\hat{\tau} < \hat{\tau}'$	Complete mediation or complete confounding	$\hat{\tau} < \hat{\tau}'$	Suppression	$\hat{\tau} < \hat{\tau}'$	Mediation or confounding

Note. As described in the text, suppression may also be inconsistent mediation or negative confounding. Mediation and confounding may also be called consistent mediation or positive confounding.

third variable effect is zero, then there is neither suppression, mediation, or confounding. However, for any given sample, the estimate of the zero third variable effect will not be exact, half of the time  $|\hat{\tau}| > |\hat{\tau}'|$  suggesting mediation or confounding, and the other half of the time  $|\hat{\tau}| < |\hat{\tau}'|$  suggesting suppression. If the population direct effect is zero and the population third variable effect is nonzero, there is complete mediation or complete confounding because the entire effect of the independent variable ( $X$ ) on the dependent variable ( $Y$ ) is due to the third variable. As shown in the table, this will occur when the population third variable effect is negative and the population direct effect is zero (row 3, column 1) and also when the population third variable effect is positive and the population direct effect is zero (row 6, column 1). Cases where the population third variable effect and the population direct effect have opposite signs indicate suppression (column 2, row 6 and column 3, row 3). In the special case where the population direct effect and the population third variable effect are of the same magnitude and opposite sign, there is complete suppression, with  $\tau = 0$ . In all cases, it is possible to come to incorrect conclusions based on sample results because of sampling variability. Hence, an estimate of the variability of the third variable effect would be helpful in order to judge whether an observed third variable effect is larger than expected by chance. Methods to calculate confidence limits have been developed in the context of mediated effects. As detailed below, these methods can be used to compute confidence limits for any of the three types of third variable effects because of the

statistical equivalence of the mediation, confounding, and suppression models.

#### INTERVAL ESTIMATORS OF THIRD VARIABLE EFFECTS

A number of different analytical solutions for estimating the variance of a third variable effect  $\alpha\beta$  or  $\tau - \tau'$  are available (e.g., Aroian, 1944; Bobko & Reick, 1980; Goodman, 1960; McGuigan & Langholtz, 1988; Sobel, 1982). Simulation work has shown that the variance estimates produced by the various methods are quite similar to each other and to the true value of the variance in situations involving continuous multivariate normal data and a continuous independent variable (MacKinnon, Warsi, & Dwyer, 1995). The simplest of these variance estimates, accurate for both continuous and binary independent variables, is the first order Taylor series or the multivariate delta method solution (Sobel, 1982, 1986):

$$\sigma_{\alpha\beta}^2 = \sigma_a^2\beta^2 + \sigma_b^2\alpha^2. \tag{4}$$

The square root of this quantity provides an estimate of the standard error of the third variable effect. This standard error estimate can then be used to compute confidence limits, which provide a general method to examine sampling variability of the third variable effect. If the confidence interval includes zero, there is evidence that the third variable effect is not larger

than expected by chance.<sup>4</sup> Several researchers now advocate confidence limits because they force researchers to consider the value of an effect as well as statistical significance and the interval has a valid probability interpretation (Harlow & Mulaik, 1997; Krantz, 1999). Another variance estimator has been proposed and is the exact variance under independence or the second order Taylor series solution for the variance of the mediated effect (Aroian, 1944):

$$\sigma_{\alpha\beta}^2 = \sigma_{\alpha}^2\beta^2 + \sigma_{\beta}^2\alpha^2 + \sigma_{\alpha}^2\sigma_{\beta}^2. \quad (5)$$

The differences between the two approximations for the standard error are usually minimal, because the third term in Equation 5 is typically quite small.

### INCONSISTENT MEDIATION AND SUPPRESSION EXAMPLE

Data from an experimental evaluation of a program to prevent anabolic steroid use among high school football players (Goldberg *et al.*, 1996) provides an example of an inconsistent mediation or suppression effect. Fifteen of 31 high school football teams received an anabolic steroid prevention program and the other 16 schools formed a comparison group. There was a significant reduction in intentions to use steroids among subjects receiving the program, where the independent variable codes exposure to the program in Equation 1,  $\hat{\tau} = -.139$ ,  $\hat{\sigma}_{\tau} = .056$ , Critical Ratio =  $-2.48$ ,  $p < .05$ . The intervention program targeted several mediators, which led to a significant reduction in intention to use steroids. We focus on one mediator, reasons for using anabolic and androgenic steroids, for which there was evidence of an inconsistent mediation effect. Estimating the parameters of Equations 2 and 3 produced the following estimates and standard errors:  $\hat{\alpha} = .573$ ,  $\hat{\sigma}_{\alpha} = .105$  for the relationship between program exposure and reasons for using steroids,  $\hat{\beta} = .073$ ,  $\hat{\sigma}_{\beta} = .014$ , for the relationship between reasons for using steroids and intentions to use steroids, and  $\hat{\gamma}' = -.181$  and  $\hat{\sigma}_{\gamma'} = .056$  for the relationship between program

exposure and intentions to use steroids adjusted for the mediator. These estimates yielded a mediated effect  $\hat{\alpha}\hat{\beta} = .042$  with standard error  $\hat{\sigma}_{\alpha\beta} = .011$  (from Equation 4). Calculating a confidence interval based on these values resulted in upper and lower 95% confidence limits for the mediated effect of .020 and .064, respectively.

Note that in this example the estimate of the total effect ( $\hat{\tau} = -.139$ ) is closer to zero than the direct effect ( $\hat{\gamma}' = -.181$ ), and that the third variable ( $\hat{\alpha}\hat{\beta} = .042$ ) and direct ( $\hat{\gamma}' = -.181$ ) effects are also of opposite sign. This pattern of coefficients indicates the presence of inconsistent mediation (i.e., a suppressor effect). Although the overall effect of the program was to reduce intentions to use steroids, this particular mediational path had the opposite effect. The program appeared to increase the number of reasons to use anabolic steroids, which in turn led to increased intentions to use steroids. Fortunately there were other, larger, significant mediation effects associated with the intervention which reduced intentions to use steroids.

### CONFOUNDING EXAMPLE

A data set from the same project (Goldberg *et al.*, 1996) also provides an example of a confounder effect. One purpose of the study was to examine the relationships among measures of athletic skill. Because these variables change with age, it was important to adjust any relationships for the potential confounding effects of age. A variable coding the height of vertical leap was related to a measure of bench press performance (number of pounds lifted times number of repetitions) with  $\hat{\tau} = 8.55$  and  $\hat{\sigma}_{\tau} = 2.25$  from Equation 1. When age was included in the model, the effect was reduced ( $\hat{\tau}' = 3.42$  and  $\hat{\sigma}_{\tau'} = 2.26$  from Equation 2). The estimate of the confounder effect thus was  $\hat{\tau} - \hat{\tau}' = 8.55 - 3.42 = 5.13$ . The estimate of the standard error of this effect, calculated using the methods outlined above, was .868. The resulting upper and lower 95% confidence limits for the estimate of the confounder effect were 3.42 and 6.83, respectively, suggesting that age had a confounding effect larger than that expected by chance. This finding suggested that future studies should consider age when examining the relationship between vertical leap and bench press performance.

<sup>4</sup>A simulation study of suppression and confounding models confirmed that the statistics typically used to assess mediation are generally unbiased in assessing confounder and suppressor effects. Point estimates of mediator, suppressor, and confounder effects and their standard errors were quite accurate for sample sizes of 50 or larger for the model in Fig. 1 and multivariate normal data. Information regarding the simulation study results can be found at the following website: [www.public.asu.edu/~davidpm/ripl](http://www.public.asu.edu/~davidpm/ripl)

### CONCEPTUAL DIFFERENCES AMONG THIRD VARIABLE EFFECTS

The discussion above centered on the fact that mediation, confounding, and suppression effects can be estimated with the same statistical methods. The conceptualizations underlying the mediational and confounding hypotheses, however, are quite different. Mediation involves a distinctly causal relationship among the variables, and the direction of causation involving the mediational path is  $X \rightarrow M \rightarrow Y$ . The confounding hypothesis, on the other hand, focuses on adjustment of observed effects to examine undistorted estimates of effects. Causality is not a necessary part of a confounding hypothesis, although often the confounder and the independent variable are hypothesized to be causally related to the dependent variable. The confounder and the independent variable are specified to covary, if a relationship is specified at all. Like confounding, suppression focuses on the adjustment of the relationship between the independent and dependent variables but in the unusual case where the size of the effect actually increases when the suppressor variable is added. In the mediation framework, a suppressor model corresponds to an inconsistent mediation model where the mediated and direct effect have opposite signs.

The distinctions between mediation and confounding thus involve the directionality and causal nature of the relationships in the model. These particular aspects of model specification are not determined by statistical testing. In some cases, the nature of the variables or the study design may dictate which of the possible specifications are plausible. For example, temporal precedence among variables may make the direction of the relationship clearer, or the causal nature of the relationship may be evident in a randomized study in which the manipulation of one variable results in changes in another. In most cases, however, the researcher must rely on theory and accumulated knowledge about the phenomenon she is studying to make informed decisions about the direction and causal nature of the relationships in the proposed model.

The appropriateness of the different conceptual frameworks may also be determined by the nature of the variables studied or by the purpose of the study. Confounders are often demographic variables such as age, gender, and race that typically cannot be changed in an experimental design. Mediators are by definition capable of being changed and are often selected based on malleability. Suppressor variables

may or may not be malleable. In a randomized prevention study, mediation is the likely hypothesis, because the intervention is designed to change mediating variables that are hypothesized to be related to the outcome variable. The randomization should remove confounding effects, although confounding may be present if randomization was compromised. It is possible, however, that a mediator may actually be disadvantageous, leading to an inconsistent mediation or a suppression effect. If the purpose of an investigation is to determine whether a covariate explains an observed relationship or to obtain adjusted measures of effects, then confounding is the likely hypothesis under study. Finally, if a variable is expected to increase effects when it is included with another predictor, then suppression is the likely model.

Replication studies also provide ability to distinguish between confounding and mediation. A third variable effect detected in a single cross-sectional study may be distinguished as a mediator or a confounder in a follow-up randomized experimental study designed to change the candidate variable. If the variable is a true mediator, then changes in the dependent variable should be specific to changes in that mediator and not others. If the variable is a confounder, the manipulation should not change effects because of the lack of causal relationship between the confounder and the outcome. If a third variable effect, whether a mediator, confounder, or suppressor, is found to be statistically significant in one study, a replication study should help clarify whether the third variable is a true mediator, confounder, or suppressor or if the conclusions from the first study are a Type 1 error.

### SUMMARY

Mediation, confounding, and suppression are rarely discussed in the same research article because they represent quite different concepts. The purpose of this article was to illustrate that these seemingly different concepts are equivalent in the ordinary least squares regression model. The estimates of these third variable effects are subject to sampling variability, and as a result, in any given sample, a variable may appear to act as a mediator, confounder, or suppressor simply due to chance. To address this issue, a statistical test based on procedures to calculate confidence limits for a mediated effect was suggested for use with any of the three types of third variable ef-

fects. The statistical procedures also can be used to determine which confounders lead to a significant adjustment of an effect and whether a suppressor effect is statistically significant. New approaches to distinguish among mediation and confounding are beginning to appear. Holland (1988) and Robins and Greenland (1992) have proposed several alternatives to establishing mediation and confounding based on Rubin's causal model (Rubin, 1974) and related methods. These methods describe the conditions required to establish causal relations among variables hypothesized in the mediation model.

Although the discussion focused on adding a third variable in the study of the relationship between an independent and dependent variable, the procedures are also applicable in more complicated models with multiple mediators. In the two mediator model, for example, there can be zero, one, or two inconsistent mediation or suppressor effects. As the number of mediators increase, the different types and numbers of effects become quite complicated.

It is important to remember that the current practice of model testing, whether conducted within a mediational or a confounding context, is necessarily a disconfirmation process. A researcher tests whether the data are consistent with a given model. Although finding inconsistency suggests that the model is incorrect, finding consistency cannot be interpreted as conclusive of model accuracy (MacCallum, Wegener, Uchino, & Fabrigar, 1993; Mayer, 1996). For any given set of data, there may be a number of different possible models that fit the data equally well. The fact that mediation, confounding, and suppression models are statistically equivalent, as shown in this paper, underlines this point (Stelzl, 1986). The estimate of the third variable effect is calculated in the same manner regardless of whether the effect is causal or correlational, and regardless of the direction of the relationship. The statistical procedures provide no indication of which type of effect is being tested. That information must come from other sources.

#### ACKNOWLEDGMENTS

J.L.K. is currently at the Department of Psychology, University of Missouri-Columbia.

This research was supported by a Public Health Service grant (DA09757). We thank Stephen West for comments on this research. Portions of this paper were presented at the 1997 meeting of the American Psychological Association.

#### REFERENCES

- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice Hall.
- Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review*, *40*, 37–47.
- Aroian, L. A. (1944). The probability function of the product of two normally distributed variables. *Annals of Mathematical Statistics*, *18*, 265–271.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.
- Bobko, P., & Rieck, A. (1980). Large sample estimators for standard errors of functions of correlation coefficients. *Applied Psychological Measurement*, *4*, 385–398.
- Breslow, N. E., & Day, N. E. (1980). *Statistical methods in cancer research. Volume I—The Analysis of Case-Control Studies*. Lyon: International Agency for Research on Cancer (IARC Scientific Publications No. 32).
- Cliff, N., & Earleywine, M. (1994). *All predictors are "mediators" unless the other predictor is a "suppressor."* Unpublished manuscript.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Conger, A. J. (1974). A revised definition for suppressor variables: A guide to their identification and interpretation. *Educational Psychological Measurement*, *34*, 35–46.
- Davis, M. D. (1985). The logic of causal order. In J. L. Sullivan and R. G. Niemi (Eds.), *Sage university paper series on quantitative applications in the social sciences*. Beverly Hills, CA: Sage Publications.
- Freedman, L. S., Graubard, B. I., & Schatzkin, A. (1992). Statistical validation of intermediate endpoints for chronic diseases. *Statistics in Medicine*, *11*, 167–178.
- Goldberg, L., Elliot, D., Clarke, G. N., MacKinnon, D. P., Moe, E., Zoref, L., Green, C., Wolf, S. L., Greffrath, E., Miller, D. J., & Lapin, A. (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.
- Goodman, L. A. (1960). On the exact variance of products. *Journal of the American Statistical Association*, *55*, 708–713.
- Hamilton, D. (1987). Sometimes  $R^2 > r_{y1}^2 + r_{y2}^2$ . *American Statistician*, *41*, 129–132.
- Hansen, W. B. (1992). School-based substance abuse prevention: A review of the state-of-the-art in curriculum, 1980–1990. *Health Education Research: Theory and Practice*, *7*, 403–430.
- Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (Eds.). (1997). *What if there were no significance tests?* Mahwah, NJ: Lawrence Erlbaum.
- Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equations models. *Sociological Methodology*, *18*, 449–484.
- Horst, P. (1941). The role of predictor variables which are independent of the criterion. *Social Science Research Council Bulletin*, *48*, 431–436.
- James, L. R., & Brett, J. M. (1984). Mediators, moderators and tests for mediation. *Journal of Applied Psychology*, *69*(2), 307–321.
- Judd, C. M., & Kenny, D. A. (1981a). Process Analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, *5*(5), 602–619.
- Judd, C. M., & Kenny, D. A. (1981b). *Estimating the effects of social interventions*. New York: Cambridge University Press.
- Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing Company.

- Krantz, D. H. (1999). The null hypothesis testing controversy in psychology. *Journal of the American Statistical Association*, *44*, 1372–1381.
- Last, J. M. (1988). *A dictionary of epidemiology*. New York: Oxford University Press.
- Lazarsfeld, P. F. (1955). Interpretation of statistical relations as a research operation. In P. F. Lazarsfeld, & M. Rosenberg (Eds.), *The language of social research: A reader in the methodology of social research* (pp. 115–125). Glencoe, IL: Free Press.
- Lord, F. M., & Novick, R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, *114*, 185–199.
- MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*, *17*(2), 144–158.
- MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, *30*(1), 41–62.
- Mayer, L. (1996 Spring). *Confounding, mediation, and intermediate outcomes in prevention research*. Paper presented at the 1996 Prevention Science and Methodology Meeting, Tempe, AZ.
- McFatter, R. M. (1979). The use of structural equation models in interpreting regression equations including suppressor and enhancer variables. *Applied Psychological Measurement*, *3*, 123–135.
- McGuigan, K., & Langholtz, B. (1988). *A note on testing mediation paths using ordinary least squares regression*. Unpublished note.
- Meinert, C. L. (1986). *Clinical trials: Design, conduct, and analysis*. New York: Oxford University Press.
- Olkin, I. & Finn, J. D. (1995). Correlations redux. *Psychological Bulletin*, *118*, 155–164.
- Robins, J. M. (1989). The control of confounding by intermediate variables. *Statistics in Medicine*, *8*, 679–701.
- Robins, J. M. & Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, *3*, 143–155.
- Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*, 688–701.
- Selvin, S. (1991). *Statistical analysis of epidemiologic data*. New York: Oxford University Press.
- Sharpe, N. R., & Roberts, R. A. (1997). The relationship among sums of squares, correlation coefficients, and suppression. *The American Statistician*, *51*, 46–48.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology*, (pp. 290–312). Washington, DC: American Sociological Association.
- Sobel, M. E. (1986). Some new results on indirect effects and their standard errors in covariance structure models. In N. Tuma (Ed.), *Sociological methodology*, (pp. 159–186). Washington, DC: American Sociological Association.
- Sobel, M. E. (1990). Effect analysis and causation in linear structural equation models. *Psychometrika*, *55*(3), 495–515.
- Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causality, prediction and search*. Berlin: Springer-Verlag.
- Stelzl, I. (1986). Changing a causal hypothesis without changing the fit. Some rules for generating equivalent path models. *Multivariate Behavioral Research*, *21*, 309–331.
- Susser, M. (1973). *Causal thinking in the health sciences: Concepts and strategies of epidemiology*. New York: Oxford University Press.
- Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications, and applications. *Psychological Bulletin*, *109*, 524–536.
- Velicer, W. F. (1978). Suppressor variables and the semipartial correlation coefficient. *Educational and Psychological Measurement*, *38*, 953–958.
- Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design*. New York: McGraw-Hill.