Testing for Mediating Variables in Management Research: Concerns, Implications, and Alternative Strategies†

J. Myles Shaver*

Carlson School of Management, University of Minnesota,
321-19th Ave S., 3-365, Minneapolis, MN 55455

Due to the nature of the data employed and of the concepts tested, standard tests for mediating variables in management research can often violate an assumption on which these tests are built. As a result, estimates from standard tests of mediating variables can often lack desirable statistical properties and lead to incorrect conclusions. This article highlights when data in management research will violate assumptions underlying tests of mediating variables, suggests an alternative strategy that provides better estimates, discusses the theoretical and empirical demands of the alternative strategy, and demonstrates the magnitude to which estimates improve by using the suggested approach.

Keywords: *mediation; research methods; hypothesis test; missing variables*

The importance of developing and testing theories of mediating effects in management research is understandable. As scholars, our goal is to understand the "drivers" of particular outcomes. This often requires that we precisely hypothesize and test what influences outcomes rather than accepting that actions and outcomes are related. As a result, research into

†I appreciate helpful comments from the anonymous reviewers, the editor Daniel Feldman, Phil Bromiley, Ya-Ru Chen, Gary Dushnitsky, Dan Forbes, Susan Kimmel, Andrew King, Mike Lenox, Mary Maloney, Elizabeth Morrison, Pri Shah, Miguel Ramos, Anand Swaminathan, Andy Van de Ven, Mary Zellmer-Bruhn, and seminar participants at the University of Minnesota.

**Corresponding author. Tel.: 612 625-1824; fax: 612 626-1316.*

E-mail address: mshaver@csom.umn.edu

Journal of Management, Vol. 31 No. 3, June 2005 330-353
DOI: 10.1177/0149206304272149

mediating effects helps assess whether the relationship between two variables is direct or whether it occurs indirectly through some third (i.e., mediating) variable.

A key, and often implicit, assumption in standard tests of mediating variables is that the error terms in the set of estimated equations are uncorrelated. The nature of the data that we examine in management research is prone to violate this assumption. In particular, the existence of measurement error and missing variables—both of which are likely given the nature of the data that we employ and the nature of the concepts that we test in management research—can lead to the correlation of error terms across regression equations. If the error terms correlate and violate the underlying assumption, the consequences can be severe. The resulting coefficient estimates are biased and inconsistent (i.e., the bias and variance of the estimates do not disappear as the sample size increases). Therefore, the conclusions drawn from these tests and the implications for managers are potentially incorrect.

To address this concern, tests of mediating variables should employ system of equation estimation techniques that explicitly recognize the potential for correlation in the errors across estimated equations. The arguments in this article demonstrate how and why (a) the two-stage least squares estimator and (b) structural equation algorithms—when applied in a particular manner—will provide consistent estimates in tests of mediating variables.

To empirically demonstrate the importance of this issue, simulated data are presented to assess the consequences of violating the underlying assumption in standard tests of mediating variables. When the data are analyzed using the standard approach, incorrect conclusions are reached about mediating effects in almost every occasion. However, when the data are analyzed with a suggested approach (i.e., two-stage least squares), incorrect conclusions are reached at a level that would be attributable to chance.

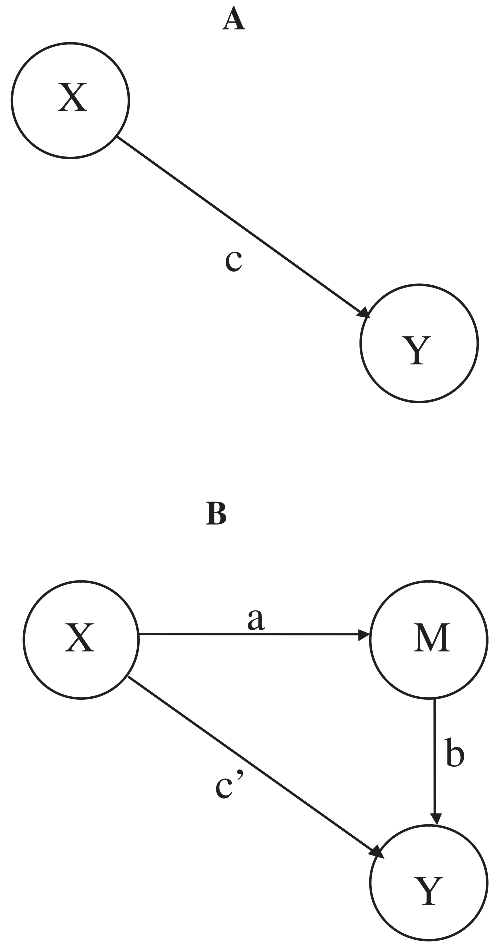
The results from the simulation analyses support the value in reconsidering tests for mediating variables in management research. Moreover, doing so will potentially be consequential because in the 2001 and 2002 volumes of the *Academy of Management Journal* and *Journal of Management*, alone, 14 studies tested for mediating variables, and all 14 are susceptible to the issue presented in this article. The appendix highlights this point by presenting an example of how the results in one of these articles change when following the approach suggested in this article.

The next section describes the standard approach to test for mediating variables. The following section describes the concern with the standard approach. It focuses on the intuition of why the assumption of uncorrelated error terms, which underlies the standard approach, can often be violated in management research. The subsequent two sections introduce a suggested approach to test for mediating variables and evaluate the suggested approach versus the standard approach.

The Standard Approach of Testing for Mediating Variables

Theories that hypothesize mediating variables take the following general form. The theory starts with an observed relationship between an independent variable (e.g., X) and a dependent variable (e.g., Y) such as indicated by Path c in Part A of Figure 1. Next, the theory argues that

Figure 1
A. Direct Relationship
B. Mediated Relationship



X might not cause *Y*. Rather, *X* affects another variable, *M*. *M*, in turn, affects *Y*. Therefore, *M* mediates the relationship between *X* and *Y*, as shown in Part B of Figure 1. A theory might or might not predict that the entire effect goes through *M*. If it does, then Path c′ will take the value zero, and the theory predicts full mediation. If the entire effect does not go through *M*, then Path c′ will be nonzero, and the theory predicts partial mediation.

To test for mediating variables, the commonly applied method requires estimating three regression equations using ordinary least squares (OLS) (e.g., Baron & Kenny, 1986). Some recent studies apply structural equation algorithms (e.g., LISREL or EQS) to test for mediating variables. Although this approach has the potential to address the issue described below, the way that these tests tend to be implemented still leaves them susceptible to the underlying concern. This point is developed further in the Suggested Approach section.

Returning to the approach described by Baron and Kenny (1986), the first step is to regress Y on X to determine if this relationship exists. This is represented by the following regression equation:

$$Y = \alpha_0 + cX + \varepsilon_0. \tag{1}$$

Consistent with the notation in Part A of Figure 1, Y is the dependent variable of interest, X is the independent variable of interest, c is a coefficient estimate of the effect of X on Y, α_0 is a coefficient estimate of the intercept, and ε_0 is the regression error term. If c is statistically different from zero, then the test of whether the effect is direct or mediated can proceed.[1] To do this, two regression equations are estimated. The first establishes whether there is a relationship between X and M by estimating the following regression equation and testing if the coefficient a is different from zero.

$$M = \alpha_1 + aX + \varepsilon_1. \tag{2}$$

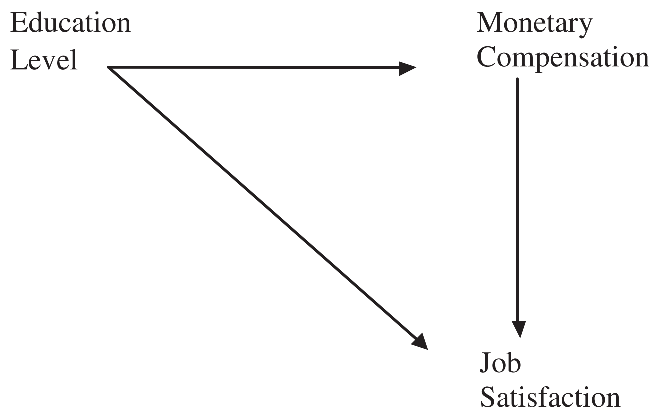
In Equation 2, α_1 is the estimate of the intercept, and ε_1 is the regression error term. If a significant relationship between X and M exists, then the final step is to assess whether X still affects Y, once controlling for the effect of M on Y. To make this assessment, the following regression equation is estimated:

$$Y = \alpha_2 + c'X + bM + \varepsilon_2. \tag{3}$$

If b is statistically significant, then given that a was statistically significant in Equation 2, the interpretation is that M mediates the relationship between X and Y. Then the estimate of c' is assessed. If c' is nonsignificant, then the interpretation is that M fully mediates the relationship between X and Y. If c' is statistically significant, then the interpretation is that M partially mediates the relationship between X and Y. This procedure of testing for mediating effects will be referred to as “the standard approach” throughout the article.

Because the article builds from these regression equations, a hypothetical example will aid in the forthcoming discussion. Assume that a positive relationship exists between a person’s level of education and his or her job satisfaction. One hypothesis would be that education increases individuals’ satisfaction with their careers through the nature of the jobs that they can obtain. However, a competing theory could hypothesize that this relationship is at least partially meditated by monetary compensation. Namely, more educated individuals earn higher monetary compensation because of their skills. Because greater monetary compensation enhances individuals’ ability to consume due to their employment, it increases job satisfaction. Figure 2 exhibits this set of relationships. As a result, education, monetary compensa-

Figure 2
Hypothetical Example



tion, and job satisfaction equate to the variables X, M, and Y, respectively, in Equations 1, 2, and 3.

Concern With the Standard Approach

There is a very important and often implicit assumption in the standard test of mediating variables. The assumption is that the regression error terms in Equations 2 and 3 (i.e., ε_1 and ε_2) are uncorrelated. Before discussing the statistical concern that stems from the correlation of error terms, the next subsection highlights why error terms might correlate given the data that we typically use and the questions that we explore in management research.

Why Would the Error Terms Correlate?

The first step to understanding why the error terms across Equations 2 and 3 might correlate is to understand what gets captured in the error term. The error term in a regression equation captures the effect of *everything* that influences the dependent variable and is not explicitly included in the set of explanatory variables. The following paragraphs consider three components of the error term: missing variables, measurement error, and truly random effects.

First, any construct that systematically affects the dependent variable yet is not explicitly included in the set of explanatory variables is called a missing variable because it is "missing" from the set of explanatory variables in the regression specification. To see this, examine Equation 2. The constant term will pick up any constant effect on M. The term aX will pick up a linear effect of X on M. These are the only two effects that are explicitly modeled in this equation. The effect of every other construct gets subsumed into ε_1. For example, if there is some other variable called U that affects the dependent variable, it gets captured by the error term. In some contexts, missing variables are also referred to as "unmeasured effects" (e.g., James & Brett, 1984) or "unobservable effects" (e.g., Shaver, 1998).

To see how a missing variable leads to the correlation of the error terms, consider the example from the previous section. Assume that there is some other systematic effect on monetary compensation that is not included in Equation 2, which is job performance (i.e., this is variable *U*). Individuals with better job performance receive greater monetary compensation compared with individuals with poor job performance. Based on the discussion in the previous paragraph, ε_1 will capture the effect of job performance on monetary compensation because it is not explicitly included in Equation 2. Therefore, ε_1 will take on greater values when job performance is good and lower values when job performance is poor. Assume also that job performance affects job satisfaction such that individuals with better job performance have higher job satisfaction because they get praised and acknowledged in nonmonetary ways. Because job performance is not explicitly included in Equation 3, ε_2 will capture the effect of job performance on job satisfaction. Therefore, ε_2 will take greater values when job performance is good and lower values when job performance is poor. Because ε_1 and ε_2 will move in tandem depending on the level of job performance, the error terms will be correlated. If the missing variable has an important effect on both variables, then the correlation (whether positive or negative) will be strong. However, if the missing variable does not have much of an effect on either variable, then the correlation (whether positive or negative) will be weaker.

Any construct that systematically affects both *M* and *Y* and is missing from the set of variables that are explicitly entered in the regression equation will cause correlation between the error terms in Equations 2 and 3. Any construct, whether previously theoretically identified or not, whether empirically measurable or not, can lead to this correlation. Therefore, even though there might not exist theories that would identify *U* and its relationship to *M* and *Y*, such an effect might still exist and correlate the errors across equations. This discussion specifies why James and Brett (1984) argued that models with serious unmeasured variables do not provide confirmatory tests of causal mediation.

The second element captured by the error term is measurement error of the dependent variable. To see this, return to Equation 2. As highlighted in the preceding discussion, ε_1 captures every element that is not explicitly included in the regression equation. Because the only effects explicitly included in this regression equation are the constant and aX, ε_1 captures the effect of measurement error in *M*. Correspondingly, because the only effects explicitly included in Equation 3 are the constant, $c'X$ and bM, ε_2 captures the effect of measurement error in *Y*. If the variables *M* and *Y* are collected by the same method or from the same respondent, then it is possible that the measurement errors of *M* and *Y* are correlated. Returning to the previous example, assume that the data on job satisfaction and monetary compensation are collected from individuals during an in-person interview. Also assume that respondents perceive there to be some social stigma with low earnings and being dissatisfied with their job. As a result, respondents have the incentive to inflate their earnings and their job satisfaction when responding to the interviewers. Therefore, the data on job satisfaction and monetary compensation are measured with error. If the measurement error with respect to an individual's report of their monetary compensation is not explicitly controlled for, then it gets captured by the error term. Likewise, if the measurement error of an individual's report of job satisfaction is not explicitly controlled for, it gets captured by the error term. As a result, systematic measurement error creates a correlation between the error terms in the two equations. Measurement error can be considered a special case of the missing variable problem. Namely, if the nature of measurement error is known, it can be controlled for and the error terms would not be correlated. When it is not controlled for, it is "missing" from the specification.

The third element captured by the error term is truly random effects. Assume that personnel records provide the data used to measure monetary compensation and a questionnaire provides the data used to measure job satisfaction. Assume also that data entry errors can occur in both the personnel data and when entering questionnaire responses. As in the previous two cases, because data entry errors are not explicitly incorporated into the regression model, they are captured by the error term. However, because such data come from different sources, the data entry errors should not be systematically related in the two variables. If the data entry errors in job satisfaction are unrelated to the data entry errors in monetary compensation, then ε_1 and ε_2 will not be correlated. Therefore, if the error terms in Equations 2 and 3 capture only random effects, it is unlikely that the error terms across equations will correlate.

To summarize, the error term in a regression equation captures truly random effects, missing variables, and systematic measurement error. If the error term captures the latter two elements, there is the likelihood that the errors between the equations in the standard tests of mediating variables will correlate. Should this occur, the data violate a key assumption in the standard test for mediating variables. As shown shortly, when these error terms correlate, the standard test for mediating variables has no desirable statistical properties (i.e., the estimates are biased and inconsistent).

Constructs and Data in Management Research

In management research, the examined constructs are often complex, multifaceted, and determined by many contextual factors (e.g., constructs like performance and satisfaction). For these reasons, it is often not possible to collect data and measure all known factors that influence the dependent variables that are examined. Moreover, all factors that influence the dependent variables might not be known. In this context, the possibility that missing variables exist and that variables are measured with error is real. For this reason, the likelihood that the error terms capture missing variables and measurement error is nontrivial. Therefore, it is likely that tests of mediating variables in management research using the standard approach violate an important assumption on which these tests are built—namely, uncorrelated regression errors.

Implications of Correlated Error Terms

As discussed previously, standard tests of mediating variables use OLS to estimate (a) the effect of X on M and (b) the effect of X on Y, controlling for the effect of M on Y (Equations 2 and 3). Because M is a dependent variable in Equation 2 and a predictor of Y in Equation 3, a system of equations exists. With such a system of equations, estimating each equation separately will provide tests with desirable properties *only* if the error terms across equations (i.e., εand ε) are uncorrelated (Kmenta, 1986). To more easily see why this occurs, substitute Equation 2 into Equation 3. This yields the following equation:

$$Y = \alpha_2 + c'X + b(\alpha_1 + aX + \varepsilon_1) + \varepsilon_2 \quad (4)$$

From Equation 4 it easily can be seen that because ε_1 partially determines M, if ε_1 and ε_2 are correlated, then M will correlate with ε_2. An important assumption of OLS is that the explanatory variables and the error term within an equation do not correlate (e.g., Kmenta, 1986: 339; Neter, Wasserman, & Kutner, 1990: 86). Violating this assumption has severe implications. When the error term and the explanatory variables correlate, the coefficient estimates are biased and inconsistent (i.e., the bias and variance of the estimates do not disappear as the sample size increases).[2] Therefore, the estimates have no desirable properties. Because the estimates of Equation 4 have no desirable statistical properties if ε_1 and ε_2 correlate, these estimates will not help inform conclusions with respect to mediating variables.

The degree to which estimates are biased is a function of the extent to which the error terms correlate. Therefore, error terms that are weakly correlated will produce smaller biases than error terms that are more strongly correlated—whether positively or negatively correlated.

To more tangibly see how correlated error terms can lead to inappropriate conclusions, return to the example of assessing if monetary compensation mediates the effect between education and job satisfaction. Assume that job performance positively affects monetary compensation and job satisfaction. In addition, assume that monetary compensation does not affect job satisfaction. Given this set of relationships, the results from estimating Equation 3 would likely result in the conclusion that monetary compensation causes job satisfaction even though this effect is not really there. The reason is that monetary compensation and job satisfaction are greater for individuals with better job performance. Because Equation 3 does not control for the effect of job performance, regressing job satisfaction on monetary compensation will show a positive effect. The intuition is that with correlated errors (ε_1 and ε_2), a statistical relationship between M and Y could be "driven" by an actual relationship between the two variables or by any other factor that affects both M and Y yet is not explicitly included in the two regression equations.

This example illustrates how the standard approach will suggest that mediation exists when it actually does not. However, the reverse condition is also possible. Missing variables or measurement errors can cause the error terms across equations to correlate in a way opposite to the true effect between M and Y. Returning to the previous example, now assume that monetary compensation positively affects job satisfaction. Furthermore, assume that overtime positively affects monetary compensation because of extra pay but negatively affects job satisfaction because of increased stress and decreased leisure time. Also assume that this is the only unmeasured effect in each equation. If Equations 2 and 3 do not control for overtime, then the error terms across Equations 2 and 3 will be negatively correlated. The negative correlation in error terms will attenuate or negate the positive effect of monetary compensation on job satisfaction when Equation 3 is estimated with OLS. This is because the effect of overtime suppresses the actual effect. Therefore, the bias in the estimates can lead to a conclusion that no effect exists when one actually does or that the effect is weaker than it actually is.

To summarize, in many management research settings, where variables are measured with error and where the constructs are complex and often contextual, it is likely that the assumption underlying standard tests of mediating variables will be violated. When this occurs, statistical estimates will have no desirable properties. Moreover, the nature of the biases of the statistical estimates will not be uniform. In some situations, the bias can result in conclusions of mediation when mediation does not actually exist. In other situations, the bias can result in conclusions that mediation does not exist when it actually exists. The following section presents an approach of testing for mediating variables that provides more meaningful estimates.

Suggested Approach

Because the possibility that the error terms in Equations 2 and 3 correlate is very real in many research settings, statistical techniques have been developed to deal with this issue, which is referred to as the simultaneous equation problem (e.g., Greene, 2000; Kmenta, 1986). In general, these techniques do not estimate the equations independently of each other but address their interdependence and estimate the equations as a system. The following discussion considers two estimation procedures: two-stage least squares (2SLS) and what is generally referred to as structural equation modeling in the management literature.

Two-Stage Least Squares

A common statistical technique that will give consistent estimates for a system of equations is 2SLS, which is an instrumental variable estimation technique. The intuition underlying 2SLS is to replace the variable M in Equation 3 with a variable that is correlated with M but not correlated with ε_2. Such a variable is called an instrument. Because the instrument is not correlated with ε_2, the estimator no longer suffers the problems described in the previous section.

A common approach is to use predicted values of M from Equation 2 as the instrument. The predicted values serve as a good instrument because if the independent variables in Equation 2 have good explanatory power of M, then the predicted values of M will be correlated with M. In addition, the predicted values of M will not be correlated with ε_1 because OLS predicted values are uncorrelated with the error term. Therefore, the predicted values of M will not correlate with ε_2 even if ε_1 and ε_2 correlate.

To summarize, in the 2SLS procedure, (a) OLS is used to estimate Equation 2, (b) predicted values of M from Equation 2 replace M in Equation 3, (c) least squares are used to estimate coefficients of Equation 3, and (d) the standard errors of these estimates are calculated. A complete discussion of 2SLS and its properties can be found in many econometric and statistics textbooks (e.g., Greene, 2000; Kmenta, 1986). Moreover, most statistical software packages provide 2SLS estimation procedures.

Baron and Kenny (1986: 1177) noted that 2SLS is an effective estimation strategy, but only when one wants to control for the possibility of feedback from the outcome to the mediator (i.e., *Y* affects *M* in addition to *M* affecting *Y*). The preceding arguments highlight why 2SLS is an effective estimation strategy in a much broader set of circumstances (i.e., even when feedback is not a concern).

Structural Equation Modeling

Alternative ways to estimate this system of equations include other techniques such as three-stage least squares or maximum likelihood. In addition to providing consistent estimates like 2SLS, these estimates will also be asymptotically efficient (a property that the 2SLS estimator does not have).[3] Common covariance-structure algorithms (e.g., LISREL and EQS) use maximum likelihood to estimate systems of equations. These algorithms, which are often referred to as structural equation models, are becoming more commonly used to test for mediating variables in management research. For this reason, the following describes under what conditions these techniques will address the concern of correlated error terms.

Common structural equation algorithms such as LISREL can estimate systems of equations where the error terms in the structural model correlate. However, this has to be explicitly modeled by the researcher. For example, in the LISREL algorithm, this occurs only if off-diagonal elements of the Ψ matrix are estimated. Although the default setting in the LISREL estimation procedure allows the off-diagonal elements of the Ψ matrix to be estimated, common practice is to "fix" (i.e., set to zero) the off-diagonal elements in order to identify the model (the issue of model identification is addressed at length in the next subsection). When researchers fix the off-diagonal elements of the Ψ matrix, they impose the assumption that the error terms across equations do not correlate. Therefore, the specification that is estimated is the same as the standard approach even though the estimation algorithm differs. Fixing the off-diagonal elements of the Ψ matrix is common practice. For example, many recent studies of mediating variables that apply structural equation techniques to test for mediating variables do not allow the error terms across equations to correlate (e.g., Allen, Shore, & Griffeth, 2003; Bacharach, Bamberger, & Sonnenstuhl, 2002; Gainey & Klaas, 2003; McAllister & Bigley, 2002; Seibert, Kraimer, & Liden, 2001).[4]

Users of the LISREL algorithm will note that the initial estimates from which the maximum likelihood estimation proceeds are often calculated by 2SLS. The reason why the initial 2SLS estimates and the final maximum likelihood estimates might substantially differ is generally the result of two factors. First, 2SLS estimates are not based on the relationships between latent variables but the relationships between one observed variable of each latent variable (i.e., a reference variable, Jöreskog & Sörbom, 1996). Second, the precise model specification might differ between the specification used to estimate the initial estimates and the maximum likelihood solution.

If the error terms are not explicitly allowed to correlate in the model specification, the use of structural equation models might mitigate the concern of correlated error terms under restrictive assumptions. This is because tests for mediation with structural equation models often test the relationships between latent variables. To the extent that (a) the estimation of latent variables significantly mitigates measurement error and (b) measurement error is the sole "driver" of correlation between the error terms in the two equations (i.e., there are no other missing variables), then these structural equations will provide consistent estimates—even if the error terms are not explicitly allowed to correlate. However, if either of these conditions does not hold, then the latent variable estimates do not adequately address the underlying issue, and the conclusions from these models are potentially incorrect.

Theoretical and Data Requirements of the Suggested Approach

A concern that arises whenever estimating a system of equations is that the system has to be identified (i.e., there have to be enough restrictions in the system to obtain numerical estimates of the parameters). For the system of Equations 2 and 3 to be identified, whether estimated by 2SLS or by a structural equation algorithm, at least one explanatory variable from Equation 2 cannot be included in Equation 3. Because Equation 2 has only one explanatory variable (i.e., X), and this variable appears in Equation 3, this system of equations is not identified, and it is not possible to obtain numerical estimates.

There are three possible ways to identify this system. First is to assume that $\varepsilon_1$and ε_2 are uncorrelated and estimate each equation separately. Of course, that is the standard approach, which the previous section noted could often be inappropriate in management research settings. It is for this reason that "fixing" the error terms across equations to zero in structural equation models will identify the system and allow them to be estimated. However, it is also for exactly this reason that most applications of structural equation models are susceptible to the concern raised in this article.

Second is to remove the term $c'X$ from Equation 3. The advantage of this solution is that the system of equations is now estimable because the removal of this explanatory variable identifies the system. However, doing so imposes the assumption that partial mediation is not possible. Namely, the effect between X and Y has to go entirely through M. Although this might be the correct specification, this is something that should be tested rather than imposed by assumption. Therefore, this solution is not entirely satisfactory.

Third is to find a variable that directly affects M but does not directly affect Y and include that variable in Equation 2. Because this additional variable does not directly affect Y, it would not be included in Equation 3, and the system would be identified. Assume that such a variable exists and label it Q. The new system of equations becomes:

$$M = \alpha_1 + aX + dQ + \varepsilon_1 \tag{5}$$

$$Y = \alpha_2 + c'X + bM + \varepsilon_2. \tag{6}$$

Because this system of equations is identified, the coefficients can be estimated by 2SLS or structural equation approaches. Therefore, a researcher can test for mediation, partial mediation, and account for the possibility that the error terms across equations correlate. Returning to the example of testing if monetary compensation mediates the effect between education and job satisfaction, such a variable could be the scarcity of people who perform the job (e.g., number of positions unfilled in a job category). Greater scarcity of people to perform the job should increase monetary compensation; however, it is likely that scarcity of people to perform the job will not be directly related to job satisfaction.

This third solution to the identification problem poses three demands on the researcher. First, they must have a theoretical reason to expect that a variable with these characteristics exists. Namely, theory should inform why a particular construct would directly affect the mediating variable but not directly affect the outcome variable. Second, they must gather data on this variable. Third, it is important that this variable explains a nontrivial amount of the variance of *M*; otherwise, the estimates will be biased toward the OLS estimates as the incremental R^2 of this variable approaches zero (Bound, Jaeger, & Baker, 1995). Provided that these three demands can be satisfied, this is the preferred way to identify the system because it still allows the test of partial versus full mediation.

Estimating a system of equations, as suggested, allows for the possibility that the error terms across equations correlate, rather than assuming that they do not. It does not impose a correlation on the error terms across equations. Therefore, employing the suggested approach does not preclude the possibility that the errors will be found not to correlate. If the errors across equations do not correlate, then tests that estimate mediation as a system of equations will provide equivalent conclusions to the standard approach.[5] Nevertheless, if the errors correlate, the conclusions drawn from the standard approach will often be incorrect.

Simulation

The previous discussion describes the statistical theory that calls into question the standard approach of testing for mediating variables due to the nature of the data that we employ in management research. However, the importance of this concern is often not appreciated unless it is connected to data.

This section presents results from simulated data that violate the assumption underlying the standard approach. Simulated data are presented because with these data the underlying structure is known with certainty. This, in turn, allows accurate assessment of the difference in results between the standard and suggested approach. When the standard approach is applied on these data, significant biases appear and incorrect conclusions are drawn. Yet, employing the suggested approach (i.e., 2SLS) largely eliminates these problems.

This analysis consists of 100 samples each with 100 observations. The data are such that *X* affects *Y* and *X* affects *M*. However, *M* does not affect *Y*. Therefore, there is no mediating effect in these data. The data structure is more easily understood by returning to the example of examining if monetary compensation mediates the effect of education on job satisfaction. This data structure is consistent with the situation where education affects monetary compensation and job satisfaction. Monetary compensation does not affect job satisfaction. Namely, there is no mediating effect. However, there is a missing variable, such as job performance, that enhances job satisfaction and monetary compensation. Figure 3 diagrammatically presents the data structure.

Figure 3
Underlying Structure of the Simulated Data

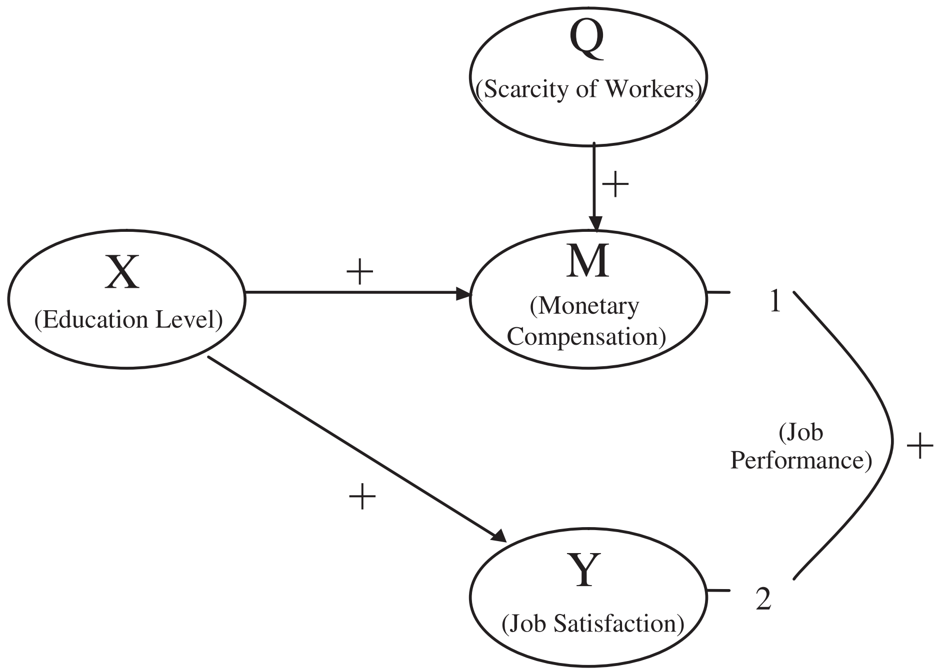


Table 1, Part A presents the precise definitions of the variables. *X* is randomly drawn. *M* is a function of a constant, *X*, the variable *U*, the variable *Q*, and a random component. *U* is included to generate correlation in the error terms of the equations. *Q* is included as a determinant of *M* in order to identify the equation for the reasons presented in the preceding section. *Y* is a function of a constant, *X*, the variable *U*, and a random component. Because *U* will not be controlled for in the following regression equations, its effect will be captured by the error terms. Therefore, the error terms will correlate across equations due to the missing variable *U*. The positive correlation should result in OLS estimates that indicate a positive relationship between *M* and *Y*, even though one does not actually exist.

The choice of structuring the simulation so that the correlation between error terms causes a spurious finding is arbitrary. It is equally possible to structure the simulation so that correlation between error terms suppresses an actual relationship. Because the motivation of this section is to provide a numerical example of the issue discussed in this article, only one of these two possibilities is presented for parsimonious versus substantive reasons.

The simulation was performed using the LIMDEP statistical package. The random variables were created using LIMDEP's random variable function. The regression analyses were performed using the "regress" procedure, and the 2SLS analyses were performed using the "2sls" procedure. The program file is available on request.

Table 1
Data Structure for the Simulation

A. 100 samples of 100 observations were generated using the following data structure

Variable	Definition
X	Random variable drawn from a normal distribution with mean = 0 and standard deviation = 1
M	$1 + X + Q + U$ + random variable drawn from a normal distribution with mean = 0 and standard deviation = 1
Y	$-1 + X + U$ + random variable drawn from a normal distribution with mean = 0 and standard deviation = 1 *Note:* Y is not a function of M (i.e., M does not mediate the effect of X on Y).
U	Random variable drawn from a normal distribution with mean = 0 and standard deviation = 1
Q	Random variable drawn from a normal distribution with mean = 0 and standard deviation = 1

B. Average correlations among the variables of interest in the 100 samples:

	X	M	Y
X	1.00		
M	.51	1.00	
Y	.58	.57	1.00

Table 1, Part B shows the average correlation among the variables X, M, and Y. Note that even though M does not determine Y and Y does not determine M, these variables are highly correlated ($r = .57$) due, in part, to the missing variable U.

Following the standard approach of testing for mediating variables, the first step is to establish a relationship between X and Y. Table 2, Part A summarizes the results from the OLS estimates of regressing Y on X for the 100 samples. The left column presents which coefficient estimate is being examined. The second column is the average value the coefficient estimate takes across the 100 samples. Below this number in parentheses is the actual value that was used to generate the data. The next five columns summarize the level of statistical significance that the coefficient estimates exhibit across the 100 samples. The third column indicates the number of samples where the t statistic of the coefficient estimate was less than –2.326, which would indicate $p < .01$ using a one-tailed test. The fourth column indicates the number of samples where the t statistic of the coefficient estimate was greater than –2.326 yet less than –1.645, which would indicate $.01 < p < .05$ using a one-tailed test. The fifth column indicates the number of samples where the t statistic of the coefficient estimate was between –1.645 and 1.645, which would indicate no relationship using the 5% cutoff. The sixth column indicates the number of samples where the t statistic of the coefficient estimate was greater than 1.645 and less than 2.326, which would indicate $.01 < p < .05$ using a one-tailed test. Finally, the seventh column indicates the number of samples where the t statistic of the coefficient estimate was greater than 2.326, which would indicate $p < .01$ using a one-tailed test.

In all 100 samples, the intercept was negative and highly significant. Moreover, the average estimate of the intercept (–1.02) is close to the true value of the intercept (–1.00). In addition, the effect of X on Y is significant at the 1% level in all samples. The average value of the estimate of c is 0.99, which is very close to the true value of 1.00.

Table 2
Simulation Results (100 samples of 100 observations)

A. Ordinary least squares (OLS) estimates of $Y = \alpha_0 + cX + \varepsilon_0$

	Average Estimate (Actual Value)	$t < -2.326$	$-2.326 < t < -1.645$	$-1.645 < t < 1.645$	$1.645 < t < 2.326$	$t > 2.326$
α_0	−1.02 (−1.00)	1.00[a]				
c	0.99 (1.00)					100

B. OLS estimates of $M = \alpha_1 + aX + dQ + \varepsilon_1$

	Average Estimate (Actual Value)	$t < -2.326$	$-2.326 < t < -1.645$	$-1.645 < t < 1.645$	$1.645 < t < 2.326$	$t > 2.326$
α_1	0.99 (1.00)					100
a	1.01 (1.00)					100
d	0.98 (1.00)					100

C. OLS estimates of $Y = \alpha_2 + c'X + bM + \varepsilon_2$

	Average Estimate (Actual Value)	$t < -2.326$	$-2.326 < t < -1.645$	$-1.645 < t < 1.645$	$1.645 < t < 2.326$	$t > 2.326$
α_2	−1.33 (−1.00)	100				
c'	0.66 (1.00)				1	99
b	0.32 (0.00)			1	*3*	*96*

D. Two-stage least squares estimates of $M = \alpha_1 + aX + dQ + \varepsilon_1$
$Y = \alpha_2 + c'X + bM + \varepsilon_2$

	Average Estimate (Actual Value)	$t < -2.326$	$-2.326 < t < -1.645$	$-1.645 < t < 1.645$	$1.645 < t < 2.326$	$t > 2.326$
α_2	–0.98 (–1.00)	98	2			
c'	1.02 (1.00)					100
b	–0.04 (0.00)		*2*	93	*3*	*2*

a. Each cell reflects the total number of samples where the t-value of the coefficient estimate in the left-most column falls within the range indicated on the top row. Numbers in italics reflect estimates that lead to conclusions that are *not* consistent with the underlying structure of the data.

The second step in the standard approach is to show that *X* affects the mediating variable *M*. Table 2, Part B presents the estimates of regressing *M* on *X* and *Q*. The coefficient estimates lead to conclusions that are consistent with the structure of the data. The estimates of the constant, the effect of *Q* on *M*, and the effect of *X* on *M* are always significant at the 1% level. The average estimates of the constant, the effect of *Q* on *M*, and the effect of *X* on *M* are 0.99, 1.01, and 0.98, respectively. Again, these are very close to the true values, which in each case is 1.00.

Table 2, Part C presents the final step of the mediation analysis using the standard approach by regressing *Y* on *X* and *M*. The effect of *X* on *Y* is statistically significant in all samples as expected. Likewise, the constant is negative and significant in all samples. Keeping in mind that the data are constructed so that no relationship exists between *M* and *Y*, 99% of the samples incorrectly show a statistically significant positive relationship between *M* and *Y*, with 96% of the samples showing the positive relationship at the $p < .01$ level! Combined with the results in the previous two equations, the conclusion from this analysis would be that *M* partially mediates the relationship between *X* and *Y* in all but one sample. This, however, is incorrect because, by construction, there is no relationship between *M* and *Y* in these data. In addition to assessing the significance levels, there are important insights to be gained from assessing the average estimates for each of these coefficients. The average coefficient estimates of the constant, *X*, and *M* are –1.33, 0.66, and 0.32, respectively. However, the true values are –1.00, 1.00, and 0. It is apparent that the biases are large in this case.

Table 2, Part D presents the estimates using 2SLS to estimate the two equations. The estimates from the first stage are the same as in Table 2, Part B; therefore, the focus is on the estimates from the second stage. First, in 93 of the 100 samples the coefficient estimate of *M* on *Y* is not statistically significant, which is consistent with the structure of the data. Using a 5% cutoff for the level of significance (for a one-tailed test), this level of incorrect conclusions is not unexpected. Contrast this with the previous table where the correct conclusion would be reached in only 1 of the 100 samples. The constant and the effect of *X* on *Y* are statistically significant with the expected sign in all cases. In addition, the average coefficient estimates are much closer to the true values with the 2SLS estimates versus the standard approach. Here, the average coefficient estimates of the constant, *X*, and *M*, are –0.98, 1.02, and –0.04, respectively (the true values are –1.00, 1.00, and 0). The overall improvement in the estimates by using 2SLS versus the standard approach is striking in these data.

The biases reflected in Table 2, Part C are expected anytime there is correlation between the two error terms. Nevertheless, the magnitude of the bias is determined by the degree of correlation between the error terms. In data where the correlations are weaker, the biases are smaller, and in data where the correlations are stronger, the biases are larger. Also, because the tests are summarized by presenting whether the coefficient estimates statistically differ from zero, the distribution of *t*-values can change by altering the magnitude of the underlying effect. For instance, if the underlying effect of *X* on *Y* is made smaller, then the average *t*-value decreases. Likewise, increasing the magnitude of the effect of *X* on *Y* will increase the average *t*-values. However, this tends to shift the average *t*-value in both Part C and Part D of Table 2. It does not alter the overall conclusions. Therefore, provided that the errors across regression equations correlate, the conclusions presented in the table will be exhibited for a wide range of parameters.

In summary, the simulation results confirm the previous sections' arguments. The standard approach of testing for mediation separately estimates regression models where the mediating variable and the outcome variable serve as dependent variables. Estimating separate regression models leads to incorrect conclusions if the error terms across these equations correlate. However, estimating the equations as a system, as suggested in this article, leads to estimates that are much more consistent with the underlying structure of the data.

To further highlight how employing the suggested approach can result in different conclusions, the appendix presents an analysis that reassesses the findings in a recently published article (McAllister & Bigley, 2002). The conclusions change markedly when the suggested approach is implemented.

Conclusion

The standard test of mediating variables employed in management research requires an assumption that is easily violated in many management research settings. Therefore, the standard approach will often be inappropriate because the results from this approach are potentially biased and the conclusions that we draw from the results incorrect.

The goals of this article were fourfold. First, to make explicit the often-implicit assumption underlying tests of mediating variables; namely, that the error terms across the regression equations are not correlated. Second, to provide intuition why the nature of the data employed and the nature of the constructs examined in management research will often violate this assumption through the mechanisms of missing variables and measurement error. Third, to suggest an estimation approach that can provide consistent estimates when the standard approach is not appropriate. Fourth, to provide evidence of how severe the biases can be when the standard approach is applied to data that violate this underlying assumption and to highlight the benefits of using the suggested approach.

The recommended approach highlighted in this article is a statistical solution to the problem. One could argue that a more direct way to address this concern is to (a) measure missing variables and directly control for their effect, and (b) explicitly control for measurement error. As stated previously, a number of studies follow this latter approach by using latent variable estimation techniques. With respect to the former approach, directly measuring and controlling for factors that affect both outcome and mediating variables is desirable. However, this does not necessarily alleviate the problem because any construct—*whether previously theoretically identified or not*—can cause the correlation of error terms across equations. Therefore, even if a researcher measures and controls for all variables identified in previous empirical and theoretical studies, this does not preclude the possibility that some construct that no one has considered correlates with the error terms across equations. Moreover, measuring and controlling for all previously known effects might make research design, data collection, and interpretation so complex that it renders research on many problems infeasible. It is for this reason that the recommendation of measuring the missing variables, although valid and beneficial, is not necessarily a solution to the concern raised in this article.

Furthermore, even though one might never be able to fully specify a model for the reasons described above, the suggested approach of testing for mediating variables has the advantage that it relaxes an assumption that the standard approach imposes. Moreover, the suggested approach is built from the premise that the model is not fully specified—namely, that the error terms across equations potentially correlate.

Finally, because of the theoretical and empirical demands required to effectively estimate the suggested statistical approach, the suggested approach will sometimes not be viable. Rather than suggest that one should not proceed with empirical analyses, I recommend that the standard approach be applied with careful interpretation and with the proper caveats. The reason is that documenting empirical relationships between variables can aid researchers in assessing how prevalent such relationships are and can provide insight into what types of missing variables might "drive" observed empirical relationships. Disseminating these findings can help enhance our understanding of relationships between the constructs that we examine.

In conclusion, finding the "drivers" of many firm, group, and individual outcomes is the goal of management research. In this vein, theories of mediating effects are important elements to help uncover real versus spurious relationships in the phenomena that we study. However, to properly play this role, we must be able to effectively test such theories in order to support, refine, and improve our understanding of management outcomes. Only if we are able to accomplish this will we be able to better understand the phenomena that we study and better inform managers' actions.

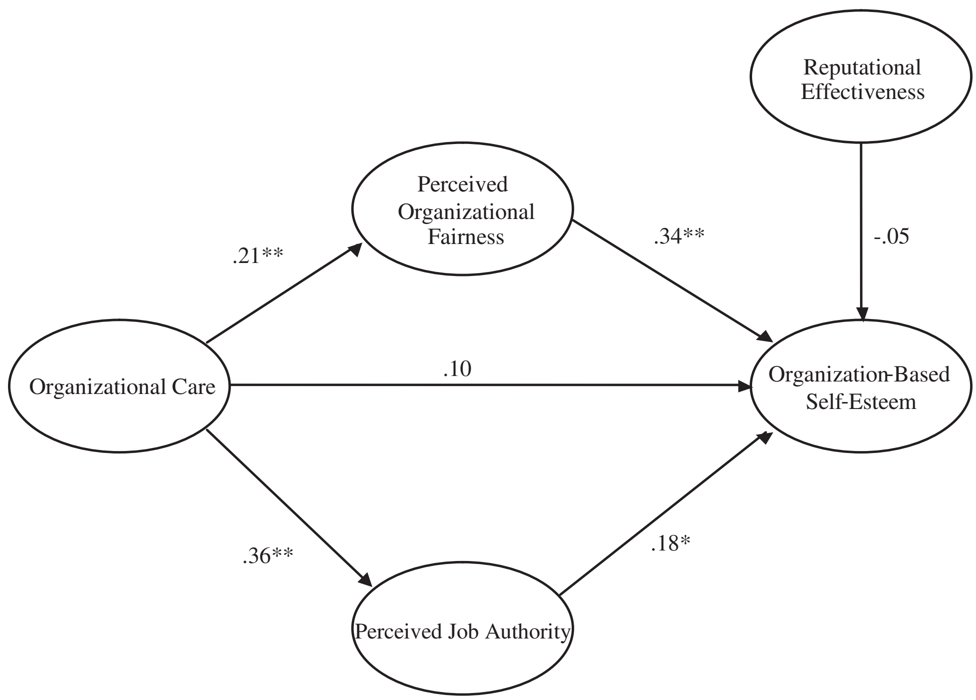
APPENDIX
Applying the Suggested Approach of Testing for Mediating Variables

In the body of the article, I argue why the standard approach of testing for mediating variables can be misleading, suggest an alternative approach, and show how the suggested approach provides better estimates than the standard approach using simulated data. In this appendix, I employ the suggested approach on data reported in a recently published article and show that the results change. As I stated in the body of the article, all 14 articles that test for mediating variables in the 2001 and 2002 editions of the *Academy of Management Journal* (*AMJ*) and the *Journal of Management* are subject to the concern that I raise because they either (a) use ordinary least squares and estimate equations separately, as in the standard approach, or (b) use structural equation models but do not allow the error terms of the mediating and outcome variables to correlate.

The article that I reexamine is McAllister and Bigley (2002), which I subsequently refer to as MB. I choose this article solely for my ability to identify the system of equations in a defendable manner, which I discuss at length in the following text. I would like to stress that, like all other articles in the 2001 and 2002 editions of *AMJ* and the *Journal of Management*, the authors assume that the error across equations does not correlate. Therefore, I would like this appendix to be interpreted not as a critique of MB but as an example of how reconsidering the standard approach of testing for mediating variables can be consequential.

MB hypothesize that perceived organizational fairness and perceived job authority mediate the effect of organization care on organization-based self-esteem. They define organization self-esteem as "an employee's evaluation of his or her personal adequacy and worthiness as an organizational member" (Gardner & Pierce, 1998: 50); organization care as "values and organization principles centered on fulfilling employees needs" (MB: p. 895); and job authority as "the amount of discretion and influence employees believe they can exercise in decisions about the work they do" (MB: p. 897). MB hypothesize that organization care affects organization fairness and that organization fairness, in turn, affects organi-

Figure A1
Replication of Findings Using the Reported Correlation Matrix



Note: $n = 186$
$*p < .05$
$**p < .01$

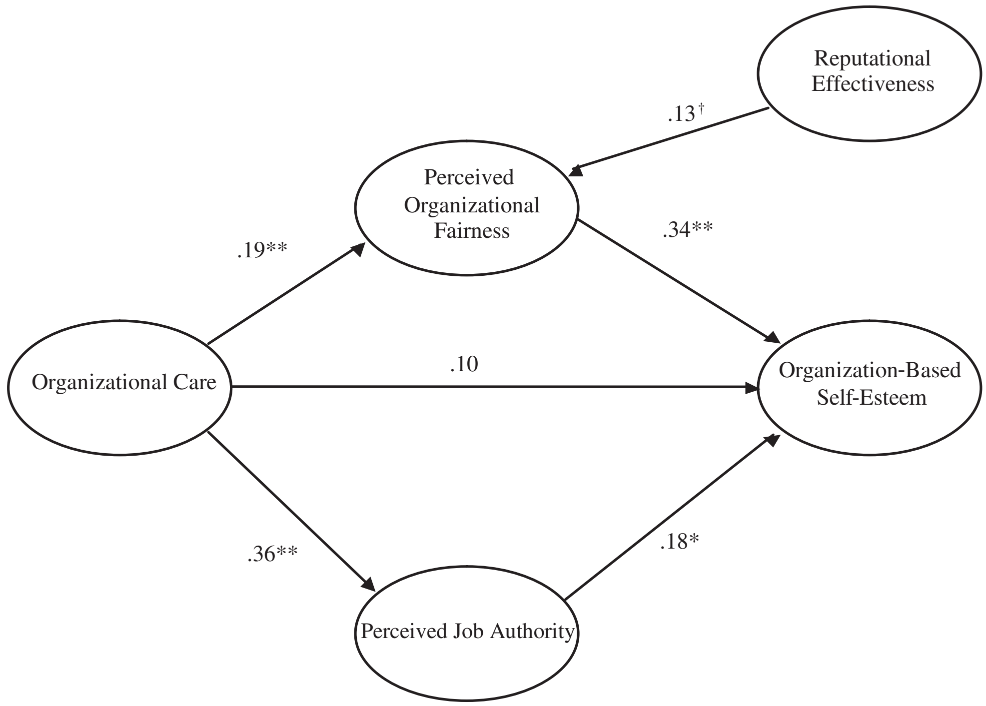
zation-based self esteem because "experienced fairness is a highly salient indicator of an individual's status in, and therefore value to, an organization" (MB: p. 896). MB also hypothesize that a caring organization empowers employees. In turn, empowerment (i.e., perceived job authority) increases organization-based self-esteem.

MB find support for both hypotheses. Using EQS to estimate a system of equations with latent variables, MB show that organization care does not directly affect organization-based self-esteem. Organization care affects perceived organizational fairness and job authority. These two constructs affect organization-based self-esteem. Although MB estimate a system of equations, they do not allow the error terms across the mediating and outcome variables to correlate as in the suggested approach. Therefore, it is susceptible to the issue that I present in the body of the article.

Before proceeding with the suggested approach, I first document that I can accurately re-create the results that MB present. I estimate their model structure using the correlation matrix in Table 1 of their article. I employ LISREL 8 to obtain the estimates; therefore, the estimation method I use is maximum likelihood rather than two-stage least squares.

Figure A1 presents my replication of MB's results. Because MB present the correlation matrix of the latent variables and not the indicator variables that form their measurement model, my estimates should

Figure A2
Model Changed to Allow for Identification of the System

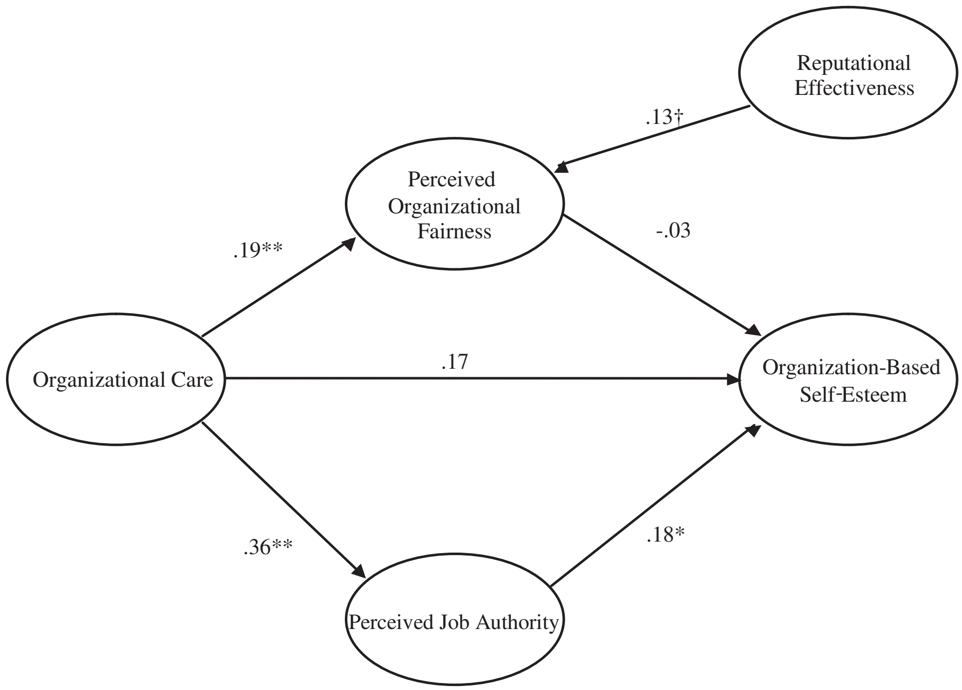


Note: n = 186. Error terms are not allowed to correlate.
†$p < .10$
*$p < .05$
**$p < .01$

not be exactly the same. Nevertheless, the estimates that I re-create are very similar. All estimates have the same threshold levels of statistical significance. With three of the six path estimates, the point estimates are the same. In the other three cases, they differ by 0.02 at the most. Given that I can effectively replicate MB's results, I can now assess if changes occur when I allow the error terms to correlate.

If I allow the error terms between the mediating and outcome variables to correlate in Figure A1, the model is not estimable because the system of equations is not identified. This is because there are no variables that predict the mediating variables and do not predict the outcome variable. However, MB include the control variable reputation effectiveness in their model. This variable measures peer assessment of how well an individual is performing his or her job. They do not find a significant effect of this variable on organization-based self-esteem. However, I develop the following hypothesis: Individuals who have positive peer assessments are more likely to consider the evaluation process fair because they can attribute the positive feedback to their efforts. Conversely, individuals who have negative peer assessments consider the process unfair and attribute the negative peer assessments to an unfair evaluation process. Therefore, I hypothesize that reputation effectiveness positively affects perceived organization fairness.

Figure A3
Identified Model With Correlated Errors Between Perceived Organizational Fairness and Organization-Based Self-Esteem



Note: $n = 186$. Estimated correlation of the errors between perceived organizational fairness and organization-based self-esteem = .37.
†$p < .10$
*$p < .05$
**$p < .01$

Moreover, I do not expect that reputation effectiveness affects organization-based self-esteem and thus do not include it as a predictor of this variable.

Figure A2 presents the results from estimating this system of equations. As expected, I find a positive effect of reputation effectiveness on perceived organization fairness, which is statistically significant at the 10% level using a two-tailed test. No other paths change threshold level of significance, and only one path changes point estimate. The effect of organization care on perceived organization fairness changes from 0.21 to 0.19. With the new structure, the underlying set of mediating relationships remains as predicted and presented by MB. Moreover, with this structure, I can allow the error terms of perceived organizational fairness and job organization-based self-esteem to correlate because that system is now identified.

Figure A3 presents the results when I allow the error terms of perceived organizational fairness and job organization-based self-esteem to correlate. The results change markedly with this specification. The error terms between perceived organizational fairness and job organization-based self-esteem positively correlate. As a result, the new estimate of the path between perceived organizational fairness and organization-based self-esteem changes from positive and significant (0.34) to negative and nonsignificant (–0.03). This suggests that some other effect that positively correlates with perceived organizational fairness and organization-based self-esteem drives the relationship between these variables. This also suggests that MB's hypothesis that perceived organizational fairness mediates the effect of organizational care and job organization–based self-esteem is not supported by the data. Moreover, the positive correlation of the error terms helps guide future theoretical development in trying to identify what variable drives the positive relationship between perceived organizational fairness and organization-based self-esteem.

To summarize, allowing the error terms to correlate across the mediating and outcome variables leads to different findings than MB report in their study. This provides a tangible example of why employing the suggested approach of testing for mediating variables can have a consequential impact on empirical findings and their interpretation.

In fairness to the authors, I had to alter the structure of their test in order to identify the system and apply the suggested approach. Although the manner in which I have done so is defendable, it might not be consistent with what the authors intended to test. Moreover, I was only able to use the correlation matrix that they presented in their published work. Therefore, the authors might possess additional data that will lead to better instruments and potentially different tests. However, that is exactly one of the points that I wish to make. If we hope to most effectively and convincingly test for mediating variables, then it is important that we consider, both theoretically and empirically, finding instruments and identifying systems of equations to test for mediating variables.

Notes

1. Stopping the analysis if c is not statistically significant assumes that the mediating effect will always be in the same direction as the effect of X on Y. This precludes testing the possibility that the effect c exists yet is offset by a counteracting mediating effect.

2. Consistency is proven by showing that the probability limit of the coefficient estimate equals the true coefficient estimate. The probability limit of the coefficient estimate equals the true coefficient estimate plus the ratio of the probability limit of the sample covariance between the explanatory variable and the error term to the probability limit of the sample variance of the explanatory variable (Kmenta, 1986: 339). Therefore, the probability limit of the coefficient estimate will only equal the true value when the probability limit of the sample covariance between the explanatory variable and the error term is zero (i.e., they are uncorrelated).

3. Asymptotic efficiency is that no other consistent estimator has smaller asymptotic variance.

4. Seibert, Kraimer, and Liden (2001) presented a model with multiple outcome variables and allow the errors of these equations to correlate. This is different than allowing the errors across equations of the mediating variables and outcome variables to correlate, which addresses the concern raised in this article.

5. From a practical perspective, the need for two-stage least squares (2SLS) can be tested using a Hausman (1978) test. This test, which is included in most statistical software packages, assesses if coefficient estimates differ between 2SLS and ordinary least squares (OLS). A statistically significant outcome rejects the null hypothesis, which is that the OLS coefficient estimates are identical to the 2SLS coefficient estimates, and indicates the need for the 2SLS estimates. A statistically insignificant outcome does not prove the null hypothesis and, therefore, cannot confirm that OLS estimates are consistent. However, a failure to reject the null hypothesis offers some justification for relying on the OLS estimates.

References

Allen, D. G., Shore, L. M., & Griffeth, R. W. 2003. The role of perceived organizational support and supportive human resource practices in the turnover process. *Journal of Management*, 29: 99-118.

Bacharach, S. B., Bamberger, P. A., & Sonnenstuhl, W. J. 2002. Driven to drink: Managerial control, work-related risk factors, and employee problem drinking. *Academy of Management Journal*, 45: 637-658.

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

Bound, J. D., Jaeger, A., & Baker, R. M. 1995. Problems with instrumental variables estimation when the correlation between the instruments and endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90: 443-450.

Gainey, T. W., & Klaas, B. S. 2003. The outsourcing of training and development: Factors impacting client satisfaction. *Journal of Management*, 29: 207-229.

Gardner, D. G., & Pierce, J. L. 1998. Self-esteem and self-efficacy within the organizational context. *Group and Organization Management*, 23: 48-70.

Greene, W. H. 2000. *Econometric analysis* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Hausman, J. 1978. Specification tests in econometrics. *Econometrica*, 46: 1251-1271.

James, L. R., & Brett, J. M. 1984. Mediators, moderators, and tests for mediation, *Journal of Applied Psychology*, 69: 307-321.

Jöreskog, K. G., & Sörbom, D. 1996. *LISREL 8 user's reference guide*. Chicago: Scientific Software International.

Kmenta, J. 1986. *Elements of econometrics* (2nd ed.). New York: Macmillan.

McAllister, D. J., & Bigley, G. A. 2002. Work context and the definition of self: How organizational care influences organization-based self-esteem. *Academy of Management Journal*, 45: 894-904.

Neter, J., Wasserman, W., & Kutner, M. H. 1990. *Applied linear models* (3rd ed.). Homewood, IL: Irwin.

Seibert, S. E., Kraimer, M. L., & Liden, R. C. 2001. A social capital theory of career success. *Academy of Management Journal*, 44: 219-237.

Shaver, J. M. 1998. Accounting for endogeneity when assessing strategy performance: Does entry mode choice affect FDI survival? *Management Science*, 44: 571-585.

Biographical Note

J. Myles Shaver (Ph.D., University of Michigan) is the Carlson School Professor of Strategic Management and Organization in the Carlson School of Management at the University of Minnesota. His research interests include the management and economics of international expansion and corporate strategy choices. He is Department Editor of Business Dynamics and Strategy at the *Journal of International Business Studies* and is on the editorial boards of the *Strategic Management Journal*, *Strategic Organization*, and the *Journal of International Business Education.*