

Improving Our Understanding of Moderation and Mediation in Strategic Management Research

Organizational Research Methods
2017, Vol. 20(4) 665-685
© The Author(s) 2016
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1094428115627498
journals.sagepub.com/home/orm



Herman Aguinis¹, Jeffrey R. Edwards²,
and Kyle J. Bradley¹

Abstract

We clarify differences among moderation, partial mediation, and full mediation and identify methodological problems related to moderation and mediation from a review of articles in *Strategic Management Journal* and *Organization Science* published from 2005 to 2014. Regarding moderation, we discuss measurement error, range restriction, and unequal sample sizes across moderator-based subgroups; insufficient statistical power; the artificial categorization of continuous variables; assumed negative consequences of correlations between product terms and its components (i.e., multicollinearity); and interpretation of first-order effects based on models excluding product terms. Regarding mediation, we discuss problems with the causal-steps procedure, inferences about mediation based on cross-sectional designs, whether a relation between the antecedent and the outcome is necessary for testing mediation, the routine inclusion of a direct path from the antecedent to the outcome, and consequences of measurement error. We also explain how integrating moderation and mediation can lead to important and useful insights for strategic management theory and practice. Finally, we offer specific and actionable recommendations for improving the appropriateness and accuracy of tests of moderation and mediation in strategic management research. Our recommendations can also be used as a checklist for editors and reviewers who evaluate manuscripts reporting tests of moderation and mediation.

Keywords

moderating effect, mediating effect, contingency, interactionism, interaction

For decades, hypotheses that involve moderation and mediation have been central to strategic management research. Moderation represents the idea that the magnitude of the effect of an antecedent (e.g., organizational structure or strategy) on firm outcomes depends on contingency

¹Department of Management and Entrepreneurship, Kelley School of Business, Indiana University, Bloomington, IN, USA

²Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Corresponding Author:

Herman Aguinis, Department of Management and Entrepreneurship, Kelley School of Business, Indiana University, 1309 E. 10th Street, Bloomington, IN 47405-1701, USA.

Email: haguinis@indiana.edu

factors, such as the uncertainty and instability of the environment and the products and services produced by the firm (e.g., Chandler, 1962; Lawrence & Lorsch, 1967; Schoonhoven, 1981; Thompson, 1967). On the other hand, mediation points to the presence of an intervening variable or mechanism that transmits the effect of an antecedent variable on an outcome (MacCorquodale & Meehl, 1948; Mathieu, DeShon, & Bergh, 2008; Ndofor, Sirmon, & He, 2011). For instance, mediation is captured by the notion that the effect of the competitive environment on firm performance is transmitted by firm strategy, such that the environment influences strategic choices that in turn affect performance (Child, 1972). In a nutshell, moderation refers to the conditions under which an effect varies in size, whereas mediation refers to underlying mechanisms and processes that connect antecedents and outcomes. Clearly, both of these pursuits are critical for advancing strategic management theory and practice.

In spite of their centrality, the assessment and interpretation of moderation and mediation are undermined by several problems. We reached this conclusion after systematically reviewing articles published in *Strategic Management Journal (SMJ)* and *Organization Science (OS)* between January 2005 and December 2014 that assessed moderation, mediation, or both. Our review of the 205 articles that assessed moderation revealed seven key problems. Overall, these demonstrated an average of 2.57 of the seven problems we identified, with only one article avoiding the problems entirely. In similar fashion, our review of the 62 articles that addressed mediation revealed six key problems, and on average, the articles exhibited 3.52 of the problems each, with none of the articles being problem-free. Accordingly, there is a need to identify and describe these problems and explain how they can be ameliorated or avoided in the future. Such treatment would contribute to “the ongoing stream of methodological inquiry in strategy research” (Wiersema & Bowen, 2009, p. 688) and benefit strategic management researchers as they pursue answers to questions that are important to field.

The goal of this article is to advance our understanding of the meaning, analysis, and interpretation of moderation and mediation in strategic management research. We do so utilizing a five-pronged approach. First, we clarify the conceptual nature and distinctions among moderation, partial mediation, and full mediation. Second, we identify key problems regarding moderation and mediation and report results of our literature review regarding their relative frequency of occurrence in articles in *SMJ* and *OS*. Third, we illustrate the detrimental impact of these issues by referring to specific substantive domains and studies. Because the problems we identified are so pervasive, we believe it would be inappropriate to identify specific articles by name. Rather, we use various research domains for illustration to engage the reader in the substantive importance of the issues involved. Fourth, we offer proposed solutions to address these problems in future research. Finally, we go beyond the more traditional treatment of moderation and mediation to explain how integrating moderation and mediation can lead to important and useful insights for strategic management theory and practice.

Moderation and Mediation: Conceptual Distinctions

A moderator variable influences the nature (e.g., magnitude and/or direction) of the effect of an antecedent on an outcome.¹ Moderation is illustrated graphically in Figure 1a, which shows that the moderator variable Z influences the path relating X to Y . When the moderator variable is categorical (e.g., industry type), the traditional data-analytic approach is subgrouping analysis, which consists of comparing correlation or regression coefficients across the various subgroups or categories (Aguinis & Pierce, 1998; Boyd, Haynes, Hitt, Bergh, & Ketchen, 2012). When the moderating effect is continuous (e.g., firm resources), studies typically rely on moderated multiple regression (Aiken & West, 1991; Cohen, 1978), which consists of creating a regression model that predicts the outcome based on a predictor X , a second predictor Z hypothesized to be a moderator, and the

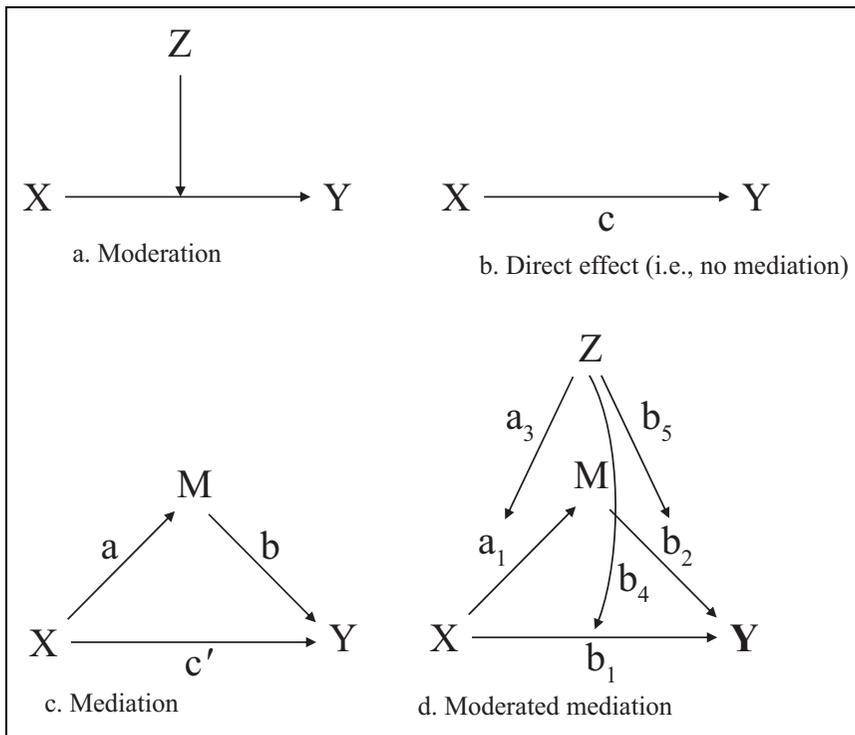


Figure 1. Graphic representation of moderation and mediation models.

product term between X and Z , which carries information on the moderating effect of Z on the X - Y relation. The regression coefficient for the XZ product term from which X and Z have been partialled out offers information on the presence as well as magnitude of the moderating effect.

A mediator variable transmits the effect of the antecedent on the outcome, either in part or whole (Baron & Kenny, 1986; MacKinnon, 2008). Figure 1b shows a no-mediation model in which there is only a direct effect of X on Y . Mediation is illustrated graphically in Figure 1c, which shows that X affects Y both directly (i.e., path c') and indirectly (i.e., the combination of paths a and b) through the mediator M . The indirect effect represents that part of the effect of X on Y that is mediated by M , with the magnitude of this effect represented by the product of the paths a and b . A full mediation model is one in which $ab \neq 0$ and $c' = 0$, whereas partial mediation exists when $ab \neq 0$ and $c' \neq 0$.

Literature Review

As mentioned earlier, we conducted a literature review of empirical articles published in *SMJ* and *OS* between January 2005 and December 2014. During this decade, *SMJ* published 794 articles, and *OS* published 717 articles, for a total of 1,511 articles. Our review included all “Research Articles” and “Research Notes and Commentaries” sections of both journals and excluded articles published in the “From the Editor” section. We used Google Scholar to search for articles that reported moderation or mediation tests. For moderation, we used the terms *interaction*, *interacting*, *moderation*, and *moderating*. This search resulted in 775 articles from *SMJ* and 696 articles from *OS*. We then manually examined these articles to determine which ones included tests of moderation using multiple regression because the majority of moderation and mediation articles use this data-analytic

approach.² The final count of articles addressing moderation was 205, which included 126 in *SMJ* and 79 in *OS*.

We conducted a similar search for mediation articles. Specifically, we used Google Scholar with the terms *indirect*, *mediate*, *mediation*, and *mediator*. This initial search resulted in a total of 315 articles published in *SMJ* and 385 articles published in *OS*. We then manually searched these articles to find those that actually conducted a mediation analysis, which yielded a final count of 62 articles, including 24 in *SMJ* and 38 in *OS*.

In the next two sections, we describe problems regarding moderation and mediation, including illustrations of the impact of these issues on substantive conclusions as well as suggested solutions for each of the problems we identified. We first address moderation, and then we turn to mediation.

Moderation: Problems and Solutions

Problem 1: Lack of Attention to Measurement Error

The most prevalent problem in strategic management studies that examine moderation concerns the effects of measurement error. Specifically, 62.44% of articles in our review did not identify measurement error as a potential problem, as evidenced by the fact that they did not mention measurement error at all. Our results specific to moderation are consistent with the finding that most articles published in *SMJ* do not report reliability estimates (Boyd, Gove, & Hitt, 2005). The reasons for this omission are unclear. It could reflect an implicit assumption that the effects of measurement error are negligible, lack of knowledge regarding the biasing effects of measurement error on parameter estimates and hypothesis tests, or prevailing norms in the domains represented by the articles.

On the surface, the lack of attention to measurement error might seem understandable for certain constructs. For example, measures of performance for public firms must go through an audit process, which leaves little room for subjectivity that might introduce measurement error (Boyd, Bergh, Ireland, & Ketchen, 2013; Dalton & Aguinis, 2013; Godfrey & Hill, 1995). However, many other constructs involve ratings of beliefs and opinions collected using self-report surveys, which are measured with error (Boyd et al., 2005). Measurement error is problematic because when independent and moderator variables are measured with error, unstandardized coefficient estimates will be biased, and this bias is particularly pronounced for moderating effects. In contrast, measurement error in outcome variables does not bias coefficient estimates, but it will attenuate estimates of explained variance, making it seem that predictors have less explanatory power than is actually the case.

Bussemeyer and Jones (1983) provided the following expression, which estimates the reliability for the product term XZ based on the reliabilities of the predictor X and moderator Z variables when both are standardized:

$$\rho_{XZ, XZ} = \frac{\rho_{XZ}^2 + \rho_{XX}\rho_{ZZ}}{\rho_{XZ}^2 + 1}. \quad (1)$$

Equation 1 indicates that when the predictor X and the moderator Z are uncorrelated (i.e., $\rho_{XZ} = 0$), the reliability of the product term is reduced to the product of the reliabilities of the predictors. For example, if the reliability of X is .70 and the reliability of Z is also .70, the resulting reliability of the product term is only .49! It seems safe to assume that few, if any, strategic management researchers would find it acceptable that 50% of the variance in a measure is random error.

There is good reason to believe that the deleterious effects of measurement error are pervasive in articles reporting tests of moderation. In fact, the vast majority of these articles report very small moderating effects across various domains, such as the moderating effect of headquarters embeddedness on the relation between subsidiary embeddedness and headquarters value-added, or the

moderating effect of a firm's resources and capabilities to deal with natural gas deregulation on the relation between managerial domain-specific experience and opportunity interpretation (i.e., ranging from threat to opportunity). In these examples, reliabilities for the product terms were not reported. Moreover, given that, when reported, reliabilities for the components are often in the .70s, it is likely that about 50% of variance in product terms is random error.

In short, tests of many moderator variable hypotheses have been undermined due to the deleterious impact of measurement error. Future research should, at a minimum, report reliability estimates for all predictors, including product components.³ Reporting reliability is particularly important for situations when a hypothesized moderating effect is not found because if reliability is low, an existing moderating effect is likely underestimated and, in some cases, might go undetected.

Problem 2: Variable Distributions Are Assumed to Include the Full Range of Possible Values

A second important problem is that samples of firms used in strategic management research usually do not represent the full range of possible scores on the variables under consideration that might exist in the population. For example, studies regarding the resource-based theory of the firm rarely include the full range of resources (Crook, Ketchen, Combs, & Todd, 2008). Similarly, firms with poor performance in the population might not be represented in the sample, which could instead consist mostly of firms with high scores on performance and related variables (Bergh et al., in press). These mechanisms lead to range restriction, meaning that the variance of variables is smaller in the sample compared to the variance in the population.

Although rarely acknowledged, range restriction has an adverse impact on tests of moderation (i.e., 34.15% of articles in our review seemed to include scores that did not span the full possible range). Specifically, Aguinis and Stone-Romero's (1997) Monte Carlo study revealed that when sample variance is less than population variance, even by what may be considered a small amount, the statistical power for detecting moderating effects is substantially diminished. For example, in a situation with a total sample size of 300 and no truncation on X scores, the statistical power to detect a medium-size moderating effect was an acceptable .81. However, when the scores were sampled from the top 80% of the distribution of the population scores, power decreased to .51. In other words, assuming that moderation exists, the accuracy of the moderating effect test is tantamount to flipping a coin. Thus, given the realistic conditions simulated by Aguinis and Stone-Romero, even a relatively mild degree of range restriction (i.e., just the bottom 20% of the distribution is truncated) can markedly decrease statistical power and threaten the validity of conclusions regarding moderating effect hypotheses. And, as with measurement error, even if a moderating effect is statistically significant, range restriction can bias the observed effect size downward. Drawing from the examples mentioned earlier, it is unlikely that variables such as a firm's resources and capabilities to deal with natural gas deregulation and managerial domain-specific experience included the full range of scores that might exist in the population of firms. Hence, most of this research has likely underestimated the true size of moderating effects.

In short, range restriction makes population-level moderating effects seem smaller than they actually are or might even render them statistically nonsignificant. Future research should attempt to capture the full range of scores of all variables involved in the analysis. When this is not feasible, and if moderating effects are small or nonsignificant, the estimated population variance should be provided to rule out range restriction as a plausible alternative explanation for the results obtained.

Problem 3: Unequal Sample Size Across Moderator-Based Categories

A third problem is that when the moderator variable is inherently categorical (e.g., industry type, firm type), the number of firms across categories is usually not equal. This issue was apparent in

20% of the articles in our review. In fact, many samples included 25% or fewer firms in one category and 75% or more firms in the other. For example, consider a study that examined type of firm ownership as a moderator of the relationship between media coverage and subsequent entry of foreign firms. The moderator “type of firm ownership” included a minority (i.e., 20%) of firms owned by individuals, while the majority (i.e., 80%) were owned by at least one firm. Under these circumstances, moderating effects are underestimated and, in many cases, can even go undetected.

This problem arises because different sample size proportions across categorical moderator variables is akin to range restriction for continuous moderator variables. Specifically, the sample variance of the dichotomous categorical moderator Z is (Aguinis, Boik, & Pierce, 2001):

$$S_Z^2 = \frac{\sum (Z_i - \bar{Z})^2}{N - 1} = \frac{Np(1 - p)}{N - 1}, \quad (2)$$

where $N = n_1$ (i.e., sample size in subgroup 1) + n_2 (i.e., sample size in subgroup 2), and $p = n_1/N$. Equation 2 shows that holding N constant, the variance of a categorical moderator is maximized when p equals .50. To illustrate, consider the following example where N is held constant at 100. If $n_1 = 20$ and $n_2 = 80$ (i.e., total $N = 100$), then the sample variance is $16/99 = .1616$. However, if $n_1 = 35$ and $n_2 = 65$ (i.e., same total $N = 100$), the sample variance increases to $22.75/99 = .2298$. If there is an even split between the two subgroups so that $n_1 = 50$ and $n_2 = 50$, the sample variance is maximized at $25/99 = .2525$.

An obvious solution to this problem is to sample a similar number of firms in each category. However, when the categorical moderator is unevenly distributed across the categories in the population, then oversampling from the smaller group improves statistical power at the cost of using a sample that might not be representative of the population. These countervailing concerns should be kept firmly in mind when investigating categorical moderator variables in strategic management.

Problem 4: Insufficient Statistical Power

Another important issue is statistical power, or the probability of finding moderating effects in a sample when they exist in the population (Aguinis, 2004; Aguinis et al., 2001). Given our previous discussion regarding measurement error, range restriction, and unequal sample sizes across categorical moderator variables, it should be apparent that the majority of research assessing moderating effects is underpowered (43.41% of articles in our review did not mention statistical power and seem underpowered given the presence of the several factors known to affect power adversely). Thus, it is not surprising that so many moderation hypotheses are not supported empirically.

To further understand the issue of statistical power, we collected additional data focusing on sample size, which, in addition to the factors mentioned earlier, is an important determinant of statistical power. Specifically, we extracted sample size information from each of the 205 articles that reported a moderating effect test. As expected based on other reviews of sample size (e.g., Shen et al., 2011), sample sizes were heavily right skewed (i.e., there are few studies including a much larger sample size than the rest). Thus, we focus on median rather than mean sample sizes. For all 205 articles combined, the median N was 227.5. The median N for *SMJ* articles was 247.5, and the median N for *OS* articles was 173. The median N of 227.5 across *SMJ* and *OS* is larger than the median N of 173 reported in *Journal of Applied Psychology* articles from 1995 to 2008 (Shen et al., 2011). However, it is still too small to yield statistical power of .80 or higher to detect the typical moderating effect size (Aguinis, Beaty, Boik, & Pierce, 2005). Thus, most tests of moderation in the studies reviewed had insufficient statistical power, which means that many moderating effects were likely to have gone undetected. Statistical power can be increased by using larger samples and conducting research in settings that control for extraneous variables (i.e., experimental or simulation-based research). In all cases, statistical power should be computed and reported to

establish whether the absence of any moderating effects can be attributed to low statistical power as opposed to the absence of a true moderating effect.

Problem 5: Artificial Dichotomization of Continuous Moderators

A fifth problem relates to the nature of the moderator variable. Many strategic management theories include hypotheses about continuous moderator variables such as environmental dynamism, firm resources, and top management team characteristics (e.g., age, tenure, compensation). When testing hypotheses about continuous moderators, a common practice is to categorize these variables into subgroups such as “high” and “low” based on whether cases fall above or below the median of the moderator variable (this problem was reported in 10.24% of articles in our review). An important shortcoming of this practice is loss of information (Cohen, 1983; MacCallum, Zhang, Preacher, & Rucker, 2002). For example, a recently published study included a measure of new ventures’ variety in the repertoire of strategic actions using ratings of 34 different actions on a 5-point scale ranging from 1 (*not a part of our strategy at all*) to 5 (*a key part of our strategy*). Although each strategic action was rated on a 5-point scale, each firm’s score was a count of the number of items rated 3 or higher. In this study, loss of information is evident because a firm that rated an action with a score of 5 and another one that rated the same action with a score of 3 were both classified as having that action as part of their strategic repertoire to the same extent. This loss of information not only undermines the interpretation of the moderator, but it also reduces the variance of the moderator variable, and the estimated moderating effects are biased downward (Aguinis, 1995). In this illustration, the moderating effect of strategic variety on the relation between a firm’s origin (i.e., whether they were created by independent entrepreneurs or established corporations) and sales growth was not statistically significant—possibly due to the artificial dichotomization of strategic variety. Furthermore, artificial dichotomization can generate nonlinear nonrandom measurement error (Maxwell & Delaney, 1993). For example, consider a moderator variable Z that ranges from 1 to 10 and dichotomized, with scores of 1 through 5 placed in one group and 6 through 10 placed in the other group. In the first group, scores of 1 and 2 will have negative measurement errors, whereas scores of 4 and 5 will have positive measurement errors. Likewise, in the second group, scores of 6 and 7 will have negative measurement errors, and scores of 9 and 10 will have positive measurement errors. This pattern of measurement errors compounds any error in the measurement of Z as a continuous variable.

In short, the practice of artificially categorizing continuous moderator variables discards information, reduces statistical power to detect moderating effects, and attenuates the size of moderating effects. Hence, this practice should be discontinued (Aguinis, 1995; Aguinis & Gottfredson, 2010).

Problem 6: Presumed Effects of Correlations Between Product Term and Its Components

A sixth issue concerns the correlation between the product term XZ and its component variables X and Z . This correlation often generates concerns about multicollinearity, as expressed in 43.90% of the articles in our review. To address this concern, researchers often center the predictor and moderator variables at their means. For example, this approach was used in a study that examined the hypothesized moderating effect of socializing on the relation between contractual complexity and procedural fairness.

Contrary to common belief, any apparent multicollinearity created by the correlation of XZ with X and Z does not cause problems for tests of moderation, provided such tests include X and Z along with XZ as predictors in the regression model. With this approach, the test for moderation does not involve XZ in its raw form, but rather the partialled XZ product, which is necessarily uncorrelated with X and Z (Cronbach, 1987; Dalal & Zickar, 2012), and the test of the partialled XZ product is

unchanged regardless of how X and Z are rescaled (Cohen, 1978). Thus, there is no need to center X and Z at their means to address this supposed “multicollinearity problem.” Nonetheless, mean-centering can facilitate the interpretation of the X and Z coefficients because these coefficients represent the slope of each variable when the other variable equals zero (Aiken & West, 1991). When X and Z are mean-centered, the coefficient for X represents its slope when Z is at its mean, and likewise, the coefficient on Z is its slope when X is at its mean.

Also related to the interpretation of first-order effects, in the presence of an XZ interaction, first-order effects can be interpreted as an average across the full range of values of the other predictor (Aiken & West, 1991). Note that the presence of a significant interaction indicates that the effect of a predictor on the outcome depends on the value of the other predictor, and consequently, an average may not be meaningful in many contexts. For example, if there is a strong and positive X - Y relation for manufacturing firms (i.e., $Z = 1$) and a strong and negative X - Y relation for service firms (i.e., $Z = 0$), the coefficient associated with X in the full model including X , Z , and XZ would lead to the misleading conclusion that, overall, there is a zero X - Y relation. These types of crossover (i.e., disordinal) interactions that involve effects in the opposite direction are not observed frequently (Rogers, 2002). Rather, noncrossover or ordinal interactions (i.e., the effect of one predictor on the outcome is in the same direction but stronger for some values of the other predictor compared to others) are more typical, and in these situations, interpreting first-order effects as an average across values of the other predictor may be informative.

In short, we recommend mean-centering for the sole purpose of facilitating the interpretation of coefficients on lower-order terms in the presence of interactions. But, it is important to recognize that results regarding interaction effects remain unchanged if predictors are centered or not.

Problem 7: Interpreting First-Order Effects Based on Models Excluding Product Terms

A seventh problem is that the coefficients on lower-order terms are often tested and interpreted prior to entering an interaction term, as evidenced in 42.93% of the studies in our review. This practice is problematic because when an interaction exists, the predictor involved in the interaction does not have a single unique effect but instead has a range of effects that vary according to the level of the moderator variable. These effects are referred to as simple slopes (Aiken & West, 1991) and can be used to interpret the form of an interaction. Because simple slopes represent a range of effects in most cases (Aguinis, 2004), it is not meaningful to hypothesize or test a single effect for a predictor when that predictor interacts with a moderator variable.

Interpreting lower-order effects from models that exclude non-zero higher-ordering effects is a pervasive problem (i.e., 42.93% of articles in our review). For example, one study hypothesized that informal control systems would enhance the performance of work units and further predicted that this relationship would be moderated by contextual factors, such as task interdependence. The initial hypothesis was tested using a regression equation that excluded the moderator variables and their products with informal control systems. These terms were subsequently added to the equation to test the moderation hypotheses, which were supported. With this approach, the results for the initial hypothesis were inconclusive because the moderator effects indicated that informal control systems did not have a simple uniform relationship with performance but instead had an array of relationships that ranged from negative to positive, depending on the levels of the three moderator variables. Situations such as this can be avoided by drawing conclusions from equations that include first-order terms along with interaction terms and drawing conclusions based on simple slopes computed and tested at meaningful levels of the moderator variables.

Summary

As a summary, Table 1 includes a brief description of each of the problems identified in our review. Although we have addressed each issue separately, Monte Carlo simulation results demonstrate that the simultaneous presence of two or more of these problems usually precludes researchers from reaching accurate conclusions about the presence of moderating effects (e.g., Aguinis, Culpepper, & Pierce, 2010; Aguinis & Stone-Romero, 1997). Thus, information in this table can be used as a resource for researchers conducting studies that involve moderation and a checklist for reviewers who evaluate manuscripts reporting tests of moderation.

Mediation: Problems and Solutions

Consistent with the broader domain of management research (Wood, Goodman, Beckmann, & Cook, 2008), the vast majority of studies in strategic management that examine mediation rely on the causal-steps procedure (Miller, Triana, Reutzler, & Certo, 2007) as set forth by Baron and Kenny (1986). This procedure can be expressed in terms of the three regression equations shown in the following:

$$Y = cX + e_Y, \quad (3)$$

$$M = aX + e_Z, \quad (4)$$

$$Y = bM + c'X + e'_Y. \quad (5)$$

In these equations, X , M , and Y are independent, mediator, and outcome variables, respectively; c , a , b , and c' , are unstandardized regression coefficients; and e_Y , e_Z , and e'_Y are residual terms (to simplify notation, we omit intercepts from the equations, as would be appropriate when M and Y are mean-centered). According to the causal-steps approach, mediation is indicated when: (a) c in Equation 3 is significant, (b) a in Equation 4 is significant, (c) b in Equation 5 is significant, and (d) c' in Equation 5 is not significant. The first condition is taken as evidence that there is a relation between X and Y to be mediated. The second and third conditions establish that the paths to and from the mediator variable M are significant. Finally, the fourth condition shows that M fully mediates the effect of X on Y . If the first three conditions are met but the fourth condition is not, then mediation is considered partial rather than complete.

Problem 1: Requiring a Significant Relation Between the Antecedent and the Outcome

Although the causal-steps approach is pervasive in studies of mediation in strategy research, it has several important drawbacks. First, requiring that c is significant in Equation 3 (as was the case in 51.61% of the articles in our review) can obscure cases in which a significant indirect effect is accompanied by a significant direct effect with the opposite sign (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). Referring to Figure 1c, the indirect effect of X on Y through M is the product ab , and the direct effect of X on Y is c' . The sum of these two effects is the total effect of X on Y , which is equivalent to c as shown in Figure 1b. Thus, c can fail to reach significance when a significant indirect effect is offset by a countervailing direct effect, leading researchers to mistakenly conclude that mediation is not present. This issue is exemplified by a recent study of the resource-based view of the firm, which invoked the first condition of the causal-steps procedure to conclude that competitive advantage did not mediate the effects of resource value on firm performance, even though resource value was related to competitive advantage, which in turn was related to firm performance.

Table 1. Summary of Problems Regarding Moderation Assessment, Detrimental Consequences for Substantive Conclusions, and Proposed Solutions.

Problems	Detrimental Consequences for Substantive Conclusions	Proposed Solution
1. Lack of attention to measurement error in tests of moderation (62.44%)	<ul style="list-style-type: none"> • Less than perfect reliability can lead to incorrectly dismissing moderating effects and to underestimating existing ones. 	<ul style="list-style-type: none"> • Report reliability estimates for all predictors (including those hypothesized to play the role of moderator variables); this practice is particularly necessary when a hypothesized moderating effect is not found.
2. Tests of moderation assume that variable distributions include the full range of possible values (34.15%)	<ul style="list-style-type: none"> • Range restriction has an adverse impact on the accuracy of substantive conclusions: Moderating effects can go undetected, and when detected, they are underestimated. 	<ul style="list-style-type: none"> • Attempt to capture the full range of population scores of all variables involved in the analysis and, if not feasible and moderating effects are not found or found to be small, information on sample and population variances should be provided to rule out range restriction as a plausible explanation for the results obtained.
3. When testing hypotheses about categorical moderators, a large total sample size is assumed to be a sufficient condition for adequate statistical power even in the presence of unequal sample size across moderator-based categories (20.00%)	<ul style="list-style-type: none"> • Even if total sample size is very large, unequal sample sizes across the moderator-based subgroups decrease statistical power and lead to an underestimation of moderating effects. 	<ul style="list-style-type: none"> • Collect data such that the number of firms within each moderator-based subgroup is similar (but keep in mind that this oversampling strategy may lead to an unrepresentative sample).
4. Statistical power is assumed to be adequate (43.41%)	<ul style="list-style-type: none"> • Statistical power is often insufficient due to small sample size, measurement error, range restriction, unequal sample size across moderator-based subgroups, and other artifacts and leads to false negative decisions regarding moderating effects. 	<ul style="list-style-type: none"> • A priori statistical power is necessary before collecting data to plan study design, and post hoc statistical power should be calculated in all cases when a moderating effect is not found to rule out the possibility that insufficient power has led to the no-moderator conclusion.
5. Moderator-based subgroups are created by categorizing continuous variables into subgroups such as “high” and “low” or categories above and below a distribution’s median value (i.e., median split procedure) (10.24%)	<ul style="list-style-type: none"> • This practice results in information loss, reduced statistical power to detect moderating effects, and a downward bias in the size of estimated moderating effects. 	<ul style="list-style-type: none"> • The practice of artificially dichotomizing or polychotomizing continuous variables should be discontinued.

(continued)

Table 1. (continued)

Problems	Detrimental Consequences for Substantive Conclusions	Proposed Solution
6. Correlations between product terms and their components are believed to be a source of bias in terms of estimating and interpreting moderating effects (43.90%)	<ul style="list-style-type: none"> Centering predictors to “reduce multicollinearity” creates additional procedures and steps that are unnecessary and create possible confusion when there is an interest in the moderating effect only and not in first-order effects. 	<ul style="list-style-type: none"> First-order predictors should be centered only if there is an interest in interpreting them in the presence of moderating effects. Also, first-order effects can be interpreted as the average effect of a predictor across the full range of values of the other predictor and their interpretation may be informative in the presence of ordinal (i.e., noncrossover) but not disordinal (i.e., crossover) interactions.
7. First-order effects are interpreted based on partial models (i.e., models not including product terms) (42.93%)	<ul style="list-style-type: none"> Because first-order coefficients are scale dependent, their interpretation in partial models leads to misleading conclusions. 	<ul style="list-style-type: none"> First-order effects should be based on centered predictors and interpreted based on full models (i.e., models including the predictor, moderator, and product terms).

Note: Values in parenthesis in the first column indicate the percentage of articles addressing moderation and published in *Strategic Management Journal* and *Organization Science* between January 2005 and December 2014 that included each problem (out of a total of 205 articles).

In short, the aforementioned study exemplifies a broader pattern exhibited by strategic management studies that apply the causal-steps procedure, which routinely excludes mediation from consideration when c in Equation 3 is not significant. This problem can be avoided by focusing on the paths that constitute the mediated effect, which together are necessary and sufficient to establish mediation (Kenny, Kashy, & Bolger, 1998).

Problem 2: Disregarding the Magnitude of the Indirect Effect

Although Baron and Kenny (1986) discussed procedures for testing the indirect effect represented by the ab product, this test is not among the conditions specified by the causal-steps approach itself. Nonetheless, this test is critical for evaluating the size of the mediating effect and comparing effects that describe alternative mediating mechanisms. As noted by Miller et al. (2007), the vast majority of strategic management studies that examine mediation do not test the mediated effect itself, and this tendency persists in recent research (i.e., 77.42% of articles in our review). For instance, a study examining decision speed as a mediator of the effects of six organizational and environmental factors on firm performance tested the individual paths of the model, but the mediating effects were not tested or compared. Consequently, this study missed an important opportunity to weigh the mediating effects of the organizational and environmental factors under consideration, which represent conceptually distinct determinants of strategic decision speed and firm performance. In the few studies that tested the mediated effect, most relied on the Sobel (1982) test and its variants, in which the product of the coefficients representing the mediated effect is divided by an estimate of its standard error and referred to a z -distribution. Unfortunately, this test is inappropriate because it rests on the assumption that the product of the coefficients is normally distributed, which is not the

case even when the coefficients themselves are normally distributed (Anderson, 1984). This assumption can be relaxed by using nonparametric testing procedures, such as percentile-based confidence intervals derived using the bootstrap (Efron & Tibshirani, 1993; Mooney & Duval, 1993; Stine, 1989). These procedures have been advocated in the methodological literature on mediation (MacKinnon, Lockwood, & Williams, 2004) and should be the method of choice for future studies of mediation in strategic management research.

Problem 3: Testing the Direct Effect as a Condition for Mediation

A third problem concerns whether c' should be tested when assessing mediation (as reported in 38.71% of articles in our review). Although this test was included in the original presentation of the causal-steps procedure (Baron & Kenny, 1986), subsequent revisions indicated that it is not required (Kenny et al., 1998). The rationale for this revision can be seen by returning to Figure 1c, which shows that mediation centers on the paths from X to M and from M to Y (i.e., a and b). The path from X to Y that bypasses M (i.e., c') need not be considered when determining whether M mediates the effect of X on Y because this path is not part of the mediated effect. Moreover, requiring that c' is not significant can cause researchers to overlook meaningful mediating processes.

To illustrate, a study of differentiation as a mediator of the effects of imitation time lag on competitiveness inferred support for mediation based on whether the effects of time lag on competitiveness was no longer significant after controlling for differentiation. By imposing this condition, several mediating effects were dismissed even though the paths involved in the effects were significant. Thus, future research should conclude that mediation exists when the indirect effect is supported, regardless of the presence or absence of a direct effect.

Problem 4: Including a Direct Effect Without Conceptual Justification

Tests of mediation that follow the causal-steps procedure routinely include a direct path from X to Y that bypasses M regardless of whether mediation is hypothesized to be complete versus partial (37.10% of articles in our review). This practice is likely due to the inclusion of X in Equation 5 and the test of c' as the fourth condition of the causal-steps procedure.

Although this practice is widespread, if the theory under investigation predicts complete mediation, then researchers should test a model that specifies complete rather than partial mediation (James, Mulaik, & Brett, 2006; LeBreton, Wu, & Bing, 2008). Referring to Figure 1, the causal-steps procedure prescribes the analysis of the model in Figure 1c regardless of whether mediation is considered partial or complete, whereas James et al. (2006) advised that researchers analyze the model in Figure 1c for partial mediation and a submodel that omits path c' for complete mediation. By following this approach, the model tested is aligned with the theory and hypotheses under consideration. Moreover, omitting path c' when complete mediation is hypothesized upholds the principle of parsimony and yields an estimate of path b that is consistent with the specified model (LeBreton et al., 2008). The consequences of omitting path c' can be assessed by testing the fit of the complete mediation model using a chi-square statistic with one degree of freedom. This test effectively compares the complete mediation model to the partial mediation model because the chi-square test captures the improvement in fit that would occur if path c' were added to the model.

In short, routinely including direct effects violates the principle of parsimony and prompts researchers to test models that are not aligned with theory. If the theory under consideration predicts mediation, future research should use the full mediation model as a baseline (i.e., $ab \neq 0$ and $c' = 0$ in Figure 1c) and test the consequences of omitting the direct effect on the fit of the model.

Problem 5: Testing Mediation With Cross-Sectional Data

Fifth, our review indicates that most strategic management studies that investigate mediation rely on cross-sectional designs (58.06% of articles in our review). However, mediated models contain causal paths that inherently involve the passage of time (Cook & Campbell, 1979; LeBreton et al., 2008; Mitchell & James, 2001), and testing these paths with cross-sectional data can produce biased estimates (Maxwell & Cole, 2007). This bias can be ameliorated by using longitudinal data in which X , M , and Y are measured sequentially on three occasions. Ideally, all three variables would be measured on each occasion, resulting in a panel model in which the hypothesized causal sequence of $X \rightarrow M \rightarrow Y$ can be compared to alternative causal flows (Finkel, 1995; Hom & Haynes, 2007). Longitudinal designs also provide a stronger basis than cross-sectional designs for drawing causal inferences, which are inherent in the interpretation of models that involve mediation (Kenny, 2008; Stone-Romero & Rosopa, 2008, 2011).

Despite their strengths, longitudinal designs do not rule out the possibility of omitted variables that can account for the relations involved in mediated models (Holland, 1988), which is a relevant source of endogeneity in strategic management research (Hamilton & Nickerson, 2003). It is encouraging that the importance of endogeneity is increasingly recognized in strategic management studies. Specifically, the author guidelines for *SMJ* note that

SMJ strongly supports research that seeks to address interesting and important questions in strategic management that involve complicated causal processes. *SMJ* recognizes that statistical analyses relevant to these questions may raise the issue of endogeneity. If relevant, authors should acknowledge this issue in submitted manuscripts, and make a good faith effort to address it. In some cases, causal inference may be impossible, but statistical correlations, especially if used to rule out some alternative hypotheses or mechanisms, may still be of interest. (“Guidelines Regarding Empirical research in *SMJ*,” 2015)

The growing attention to endogeneity is a very positive development given that, in a review of articles published in *SMJ* from 2005 and 2012, Semadeni, Withers, and Certo (2014) found that only 24 (about 4%) articles of a total of about 580 published during this period reported using two-stage least squares for testing endogeneity.

In short, mediated models contain causal paths that imply the passage of time, and testing these paths with cross-sectional data can produce biased estimates. When possible, and absent the possibility of implementing an experimental design (Eden, Stone-Romero, & Rothstein, 2015), future research should assess mediation using longitudinal data, preferably with panel models that allow the comparison of alternative causal flows.

Problem 6: Lack of Attention to Measurement Error

Finally, as with moderation, studies that examine mediation rarely address the consequences of measurement error (88.71% of articles in our review). Measurement error in X and M can bias path estimates upward or downward (Baron & Kenny, 1986).⁴ As a result, statistical tests of these paths can be either too liberal or too conservative, either of which would lead to incorrect conclusions. When mediation is examined using multiple regression, which our review indicated is usually the case (see footnote 2 and also Miller et al., 2007), measurement error is effectively disregarded, thereby introducing the aforementioned bias and its deleterious consequences.

Clearly, the best solution for addressing measurement error is to create and use more reliable measures (Aguinis & Edwards, 2014; Aguinis & Vandenberg, 2014). However, as a second best option, some of the effects of measurement error can be offset by using structural equation modeling

(SEM) with latent variables (Bollen, 1989), which has become increasingly prevalent in strategic management research (Shook, Ketchen, Hult, & Kacmar, 2004). We are certainly not suggesting that SEM is some magic cure for the many problems that can beset poor quality measures or that using SEM gives researchers an excuse to disregard fundamental measurement issues such as reliability and construct validity. Moreover, conventional applications of SEM only correct for certain sources of measurement error (DeShon, 1998), and other sources require more elaborate model specifications. Nevertheless, SEM offers important advantages over procedures that ignore measurement error completely.

Summary

The aforementioned issues concerning tests of mediation are summarized in Table 2, which gives a brief description of each problem, its consequences, and proposed solutions. As with Table 1, information in Table 2 can be used as a resource for researchers as well as a checklist for reviewers who evaluate manuscripts including mediation tests.

Integrating Moderation and Mediation

Thus far, we have discussed moderation and mediation separately. However, many theories in strategic management implicitly or explicitly combine moderation and mediation. For instance, research on contingency theory not only investigates factors that moderate the effects of firm strategy on performance but also considers the processes and mechanisms responsible for these effects (Boyd et al., 2012). Accordingly, we provide examples of research that would benefit from the integration of moderation and mediation and briefly discuss methods that can be used to accomplish this task.

Conceptually, moderation and mediation are integrated when the paths that constitute a mediated model are theorized to vary according to the level of a moderator variable. There are several domains in strategy research that warrant testing models that combine moderation and mediation. For instance, from a conceptual standpoint, the effects of competitive environment on firm performance are mediated by factors within the firm such as strategic choice. These mediated effects could depend on the resources available to the firm such that firms with greater resources at their disposal can choose from a wider array of strategic options and more readily translate these choices into gains in performance. Analogously, in the organizational learning domain, organizational failures lead to learning, but this effect could be mediated by top management team's coping mechanisms. As such, coping is an intervening mechanism for the relation between failures and learning. In turn, the strength of this mediated relation could be moderated by organizational climate such that the relation is stronger in organizations with more positive compared to more negative climates. A third illustration is the effects of strategic consensus on firm performance (Gonzalez-Benito, Aguinis, Boyd, & Suarez-Gonzalez, 2012). Consensus can refer to long-term strategic goals (i.e., consensus on objectives) or the methods that need to be implemented to achieve those goals (i.e., consensus on means). The relation between consensus on objectives and performance is likely to be moderated by environmental dynamism. At the same time, the relation between consensus and performance could be mediated by consensus on means. As these examples illustrate, we see great potential of future research assessing moderation and mediation within a given conceptual model.

A general framework for specifying and testing models that combines moderation and mediation was developed by Edwards and Lambert (2007). This framework overcomes problems with conventional methods and can be applied to a wide range of models. An illustrative model from this framework is depicted in Figure 1d, which shows that the effect of X on Y is mediated by M , and the

Table 2. Summary of Problems Regarding Mediation Assessment, Detrimental Consequences for Substantive Conclusions, and Proposed Solutions.

Problems	Detrimental Consequences for Substantive Conclusions	Proposed Solution
1. A significant relation between the antecedent and the outcome is a prerequisite for testing mediation (51.61%)	<ul style="list-style-type: none"> This assumed prerequisite can lead researchers to overlook mediation when the direct and indirect effects in a model have opposite signs. 	<ul style="list-style-type: none"> Conduct the mediation test without the precondition that the relation between the antecedent and the outcome should be significant.
2. The causal-steps procedure yields information about the magnitude of the indirect effect transmitted through the mediator (77.42%)	<ul style="list-style-type: none"> The causal-steps procedure does not involve the computation or test of the indirect effect and therefore does not indicate the magnitude of the mediating effect or allow its comparison with other effects (e.g., the indirect vs. the direct effect). 	<ul style="list-style-type: none"> Compute the size of the indirect effect by multiplying the paths to and from the mediator (i.e., paths a and b in Figure 1c) and test this product using nonparametric procedures such as the bootstrap (rather than the Sobel test that assumes normality in the distribution of product coefficients).
3. The evaluation of mediation should include a test of the direct effect that bypasses the mediator (38.71%)	<ul style="list-style-type: none"> Because the direct effect has no bearing on the presence of the indirect effect, this test can lead researchers to inappropriately dismiss mediating effects. 	<ul style="list-style-type: none"> Conclude that mediation exists when the indirect effect is supported, regardless of the presence or absence of a direct effect.
4. The causal-steps procedure routinely includes a direct path from the antecedent to the outcome regardless of whether this path is conceptually justified (37.10%)	<ul style="list-style-type: none"> Routinely including direct effects violates the principle of parsimony and prompts researchers to test models that are not aligned with theory. 	<ul style="list-style-type: none"> If the theory under consideration predicts complete mediation, then use the full mediation model as a baseline (i.e., $ab \neq 0$ and $c' = 0$ in Figure 1c) and formally test the consequences of omitting the direct effect on the fit of the model.
5. Mediation can be tested satisfactorily with cross-sectional data (58.06%)	<ul style="list-style-type: none"> Mediated models contain causal paths that imply the passage of time, and testing these paths with cross-sectional data can produce biased estimates. 	<ul style="list-style-type: none"> When possible, assess mediation using longitudinal data, preferably with panel models that allow the comparison of alternative causal flows.
6. Lack of attention to measurement error in tests of mediation (88.71%)	<ul style="list-style-type: none"> Measurement error can bias path estimates upward or downward, leading to conclusions that are unwarranted. 	<ul style="list-style-type: none"> Create and use reliable measures. As a second best option, use multiple-item measures for all constructs and analyze the data using structural equation modeling with latent variables.

Note: Values in parenthesis in the first column indicate the percentage of articles addressing mediation and published in *Strategic Management Journal* and *Organization Science* between January 2005 and December 2014 that included each problem (out of a total of 62 articles).

three paths of the model are moderated by Z . Following the notation of Edwards and Lambert (2007), the equations for this model are as follows:

$$M = a_0 + a_1X + a_2Z + a_3XZ + e_M, \quad (6)$$

$$Y = b_0 + b_1X + b_2M + b_3Z + b_4XZ + b_5MZ + e_Y. \quad (7)$$

These equations add the moderator variable Z and its products with X and M to the equations that constitute the basic three-variable mediated model (as shown in Figure 1c). The extent to which the paths of the model are moderated by Z is given by the coefficients a_3 , b_4 , and b_5 . In addition, Equation 6 can be substituted into Equation 7 to recover simple effects that indicate the extent to which the paths of the model vary across levels of Z .

Despite the conceptual relevance of integrating moderation and mediation, methods for doing so have been underutilized in strategic management studies. Out of the 1,511 articles published in *SMJ* and *OS* over the past decade, only 14 combined moderation and mediation within the same study. Moreover, 11 of these articles used methods identified as problematic by Edwards and Lambert (2007), including the application of the causal steps procedure to product terms such as XZ (Baron & Kenny, 1986). On a more positive note, the following three published articles used appropriate methods of analysis to combine moderation and mediation: Bunderson, Van der Vegt, and Sparrowe (2013); Chen, Takeuchi, and Shum (2013); and De Jong, Bijlsma-Frankema, and Cardinal (2014). Thus, these articles can be used as examples to guide future research.

Summary and Conclusions

The study of moderation and mediation is fundamental to the field of strategic management (Boyd et al., 2012; Ketchen, Boyd, & Bergh, 2008). To facilitate research in this important area of investigation, we summarized key methodological problems, documented their prevalence based on a review of articles published in *SMJ* and *OS*, explained their deleterious consequences regarding substantive conclusions, and proposed solutions for each problem. Although we focused on the strategic management literature, many of the problems we identified are also common in micro domains in management such as organizational behavior and human resource management (Aguinis et al., 2005) and other fields including information systems (Carte & Russell, 2003), accounting (Hartmann & Moers, 1999), marketing (Irwin & McClelland, 2001), education (Kromrey, & Foster-Johnson, 1998), public administration (Cascio & Aguinis, 2001), and many others (Aguinis, 2004). We also discussed how moderation and mediation can be integrated, which can stimulate new ways of theorizing.

We readily acknowledge that some of our proposed solutions might not be easy to implement because some of them require taking action before data are collected, particularly those that involve theory, design, and measurement issues rather than data-analytic issues (Aguinis & Edwards, 2014; Aguinis & Vandenberg, 2014; Eden et al., 2015; Vancouver & Carlson, 2015). Moreover, improvements in methodological practices are slow, particularly among substantive compared to researchers interested in methodological issues (Aguinis, Pierce, Bosco, & Muslin, 2009) because there is a “scientific community’s persistence in the use of particular methods” (Podsakoff & Dalton, 1987, p. 433). Nevertheless, we add our voice to those of others (e.g., Bettis, Ethiraj, Gambardella, Helfat, & Mitchel, in press; Bettis, Gambardella, Helfat, & Mitchell, 2014; Wiersema & Bowen, 2009) who have shown that the potential for misusing methods is not uncommon and impedes theoretical progress. We hope our article will serve as a useful resource for current and future scholars as well as journal editors and reviewers. Moreover, our hope is that this accumulating body of methodological work will be used for training future generations of strategic management researchers because, as noted by philosopher and poet Jorge Santayana, those who cannot remember the past are condemned to repeat it.

Acknowledgments

The first and second authors contributed equally to this research. We thank James LeBreton, Brian Boyd, Jose Cortina, Jeff Covin, Dan Dalton, Dean Shepherd, and Margarethe Wiersema for highly constructive and helpful feedback that allowed us to improve our manuscript substantially.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Although we used the term *effect* throughout the manuscript and its usage implies causality, moderation and mediation are usually assessed in strategic management studies using non-experimental (i.e., observational) research designs, which preclude strong inference about causal relations (Aguinis & Edwards, 2014; Eden, Stone-Romero, & Rothstein, 2015; Spencer, Zanna, & Fong, 2005; Vancouver & Carlson, 2015). However, we use the term *effect* as a matter of convenience.
2. Although other data-analytic approaches were also used to assess moderation or mediation, multiple regression is by far the most popular. For example, only 3 articles used structural equation modeling to assess moderation (15 for mediation), only 33 articles used generalized least squares regression to assess moderation (6 for mediation), and only 46 articles used logistic regression to assess moderation (0 for mediation). Thus, we focus on moderation and mediation tests involving multiple regression.
3. Note that the reliability of the product term is scale dependent (Bohrnstedt & Marwell, 1978). Thus, it is not possible to obtain an accurate reliability estimate unless the components are measured with ratio scales (i.e., true zero point), which is the case for some (i.e., measures of financial-based constructs such as return on investment and Tobin's Q) but not the majority of variables in strategic management studies. As noted by Edwards (2008), ρ_{XZ} depends on the scales of X and Z , and this is why Bohrnstedt and Marwell (1978) cautioned against estimating ρ_{XZ} when X and Z are measured on interval scales because the estimated value of ρ_{XZ} is as arbitrary as the origins of X and Z .
4. As noted by Baron and Kenny (1986), "The presence of measurement error in the mediator tends to produce an underestimate of the effect of the mediator and an overestimate of the effect of the independent variable on the dependent variable when all coefficients are positive" (p. 1177). As their source, Baron and Kenny cited Judd and Kenny (1981), who explained that measurement error in the mediator M will bias its coefficient downward but can bias the coefficient on the predictor X upward or downward. The direction of this bias depends on the correlations among X , M , and the outcome Y . Specifically, when r_{XY} is less than r_{XM}/r_{MY} , the coefficient on X will be biased upward, whereas when r_{XY} is greater than r_{XM}/r_{MY} , the coefficient on X will be biased downward. In either case, the coefficient on M will be biased downward. The effects of measurement error in X follow an analogous pattern, such that the coefficient on X is biased downward and the coefficient on M is biased upward or downward, depending on whether r_{MY} is less than or greater than r_{XM}/r_{XY} , respectively. This phenomenon is also discussed in texts on regression analysis (e.g., Cohen, Cohen, West, & Aiken, 2003, pp. 121-124) and structural equation modeling (e.g., Bollen, 1989, chapter 5).

References

- Aguinis, H. (1995). Statistical power problems with moderated multiple regression in management research. *Journal of Management*, *21*, 1141-1158.
- Aguinis, H. (2004). *Regression analysis for categorical moderators*. New York, NY: Guilford.
- Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology*, *90*, 94-107.
- Aguinis, H., Boik, R. J., & Pierce, C. A. (2001). A generalized solution for approximating the power to detect effects of categorical moderator variables using multiple regression. *Organizational Research Methods*, *4*, 291-323.

- Aguinis, H., Culpepper, S. A., & Pierce, C. A. (2010). Revival of test bias research in preemployment testing. *Journal of Applied Psychology, 95*, 648-680.
- Aguinis, H., & Edwards, J. R. (2014). Methodological wishes for the next decade and how to make wishes come true. *Journal of Management Studies, 51*, 143-174.
- Aguinis, H., & Gottfredson, R. K. (2010). Best-practice recommendations for estimating interaction effects using moderated multiple regression. *Journal of Organizational Behavior, 31*, 776-786.
- Aguinis, H., & Pierce, C. A. (1998). Heterogeneity of error variance and the assessment of moderating effects of categorical variables: A conceptual review. *Organizational Research Methods, 1*, 296-314.
- Aguinis, H., Pierce, C. A., Bosco, F. A., & Muslin, I. S. (2009). First decade of *Organizational Research Methods*: Trends in design, measurement, and data-analysis topics. *Organizational Research Methods, 12*, 69-112.
- Aguinis, H., & Stone-Romero, E. F. (1997). Methodological artifacts in moderated multiple regression and their effects on statistical power. *Journal of Applied Psychology, 82*, 192-206.
- Aguinis, H., & Vandenberg, R. J. (2014). An ounce of prevention is worth a pound of cure: Improving research quality before data collection. *Annual Review of Organizational Psychology and Organizational Behavior, 1*, 569-595.
- Aiken, L. A., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Anderson, T. W. (1984). *An introduction to multivariate statistical analysis* (2nd ed.). New York, NY: Wiley.
- 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.
- Bergh, D. D., Aguinis, H., Heavey, C., Ketchen, D. J., Boyd, B. K., Su, P., . . . Joo, H. (in press). Using meta-analytic structural equation modeling to advance strategic management research: Guidelines and an empirical illustration via the strategic leadership-performance relationship. *Strategic Management Journal*. doi:10.1002/smj.2338
- Bettis, R. A., Ethiraj, S., Gambardella, A., Helfat, C., & Mitchel, W. (in press). Creating repeatable cumulative knowledge in strategic management. *Strategic Management Journal*. doi:10.1002/smj.2477
- Bettis, R., Gambardella, A., Helfat, C., & Mitchell, W. (2014). Quantitative empirical analysis in strategic management. *Strategic Management Journal, 35*, 949-953.
- Bohrstedt, G. W., & Marwell, G. (1978). The reliability of products of two random variables. In K. F. Schuessler (Ed.), *Sociological methodology* (pp. 254-273). San Francisco, CA: Jossey-Bass.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.
- Boyd, B. K., Bergh, D. D., Ireland, R. D., & Ketchen, D. J. (2013). Constructs in strategic management. *Organizational Research Methods, 16*, 3-14.
- Boyd, B. K., Gove, S., & Hitt, M. A. (2005). Construct measurement in strategic management research: Illusion or reality? *Strategic Management Journal, 26*, 239-257.
- Boyd, B. K., Haynes, K. T., Hitt, M. A., Bergh, D. D., & Ketchen, D. J. (2012). Contingency hypotheses in strategic management research: Use, disuse, or misuse? *Journal of Management, 38*, 278-313.
- Bunderson, J. S., Van der Vegt, G. S., & Sparrowe, R. T. (2013). Status inertia and member replacement in role-differentiated teams. *Organization Science, 25*, 57-72.
- Busemeyer, J. R., & Jones, L. E. (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin, 93*, 549-562.
- Carte, T. A., & Russell, C. J. (2003). In pursuit of moderation: Nine common errors and their solutions. *MIS Quarterly, 27*, 479-501.
- Cascio, W. F., & Aguinis, H. (2001). The Federal Uniform Guidelines on Employee Selection Procedures (1978): An update on selected issues. *Review of Public Personnel Administration, 21*, 200-218.
- Chandler, A. D. (1962). *Strategy and structure in resources, firms, and strategies*. Cambridge, MA: MIT Press.
- Chen, Z., Takeuchi, R., & Shum, C. I. (2013). A social information processing perspective of coworker influence on a focal employee. *Organization Science, 24*, 1618-1639.

- Child, J. (1972). Organizational structure, environment and performance: The role of strategic choice. *Sociology*, 6, 1-22.
- Cohen, J. (1978). Partialled products *are* interactions; partialled powers *are* curve components. *Psychological Bulletin*, 85, 858-866.
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, 7, 249-253.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple correlation/regression analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin Company.
- Cronbach, L. J. (1987). Statistical tests for moderator variables: Flaws in analyses recently proposed. *Psychological Bulletin*, 102, 414-417.
- Crook, T. R., Ketchen, D. J., Combs, J. G., & Todd, S. Y. (2008). Strategic resources and performance: A meta-analysis. *Strategic Management Journal*, 29, 1141-1154.
- Dalal, D. K., & Zickar, M. J. (2012). Some common myths about centering predictor variables in moderated multiple regression and polynomial regression. *Organizational Research Methods*, 15, 339-362.
- Dalton, D. R., & Aguinis, H. (2013). Measurement malaise in strategic management studies: The case of corporate governance research. *Organizational Research Methods*, 16, 88-99.
- De Jong, B. A., Bijlsma-Frankema, K. M., & Cardinal, L. B. (2014). Stronger than the sum of its parts? The performance implications of peer control combinations in teams. *Organization Science*, 25, 1703-1721.
- DeShon, R. P. (1998). A cautionary note on measurement error corrections in structural equation models. *Psychological Methods*, 3, 412-423.
- Eden, D., Stone-Romero, E. F., & Rothstein, H. R. (2015). Synthesizing results of multiple randomized experiments to establish causality in mediation testing. *Human Resource Management Review*, 25, 342-351.
- Edwards, J. R. (2008). Seven deadly myths of testing moderation in organizational research. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Received doctrine, verity, and fable in the organizational and social sciences* (pp. 145-166). New York, NY: Routledge.
- Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods*, 12, 1-22.
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.
- Finkel, S. E. (1995). *Causal analysis with panel data*. Newbury Park, CA: Sage.
- Godfrey, P. C., & Hill, C. W. L. (1995). The problem of unobservables in strategic management research. *Strategic Management Journal*, 16, 513-533.
- Gonzalez-Benito, J., Aguinis, H., Boyd, B. K., & Suarez-Gonzalez, I. (2012). Coming to consensus on strategic consensus: A mediated moderation model of consensus and performance. *Journal of Management*, 38, 1685-1714.
- Guidelines regarding empirical research in *SMJ*. (2015). Retrieved from <http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291097-0266/homepage/ForAuthors.html>.
- Hamilton, B. H., & Nickerson, J. A. (2003). Correcting for endogeneity in strategic management research. *Strategic Organization*, 1, 51-78.
- Hartmann, F. G. H., & Moers, F. (1999). Testing contingency hypotheses in budgetary research: An evaluation of the use of moderated regression analysis. *Accounting, Organizations, and Society*, 24, 291-315.
- Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equations models. In C. C. Clogg (Ed.), *Sociological methodology* (pp. 449-484). Washington, DC: American Sociological Association.
- Hom, P., & Haynes, K. T. (2007). Applying advanced panel methods to strategic management research: A tutorial. In D. J. Ketchen & D. D. Bergh (Eds.), *Research methodology in strategy and management* (Vol. 4, pp. 193-272). Bingley, UK: Emerald.
- Irwin, J. R., & McClelland, G. H. (2001). Misleading heuristics and moderated multiple regression models. *Journal of Marketing Research*, 38, 100-109.

- James, L. R., Mulaik, S. A., & Brett, J. M. (2006). A tale of two methods. *Organizational Research Methods, 9*, 233-244.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5*, 602-619.
- Kenny, D. A. (2008). Reflections on mediation. *Organizational Research Methods, 11*, 353-358.
- Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. *The Handbook of Social Psychology, 1*, 233-265.
- Ketchen, D. J., Boyd, B. K., & Bergh, D. D. (2008). Research methodology in strategic management past accomplishments and future challenges. *Organizational Research Methods, 11*, 643-658.
- Kromrey, J. D., & Foster-Johnson, L. (1998). Mean centering in moderated multiple regression: Much ado about nothing. *Educational and Psychological Measurement, 58*, 42-67.
- Lawrence, P. R., & Lorsch, J. W. (1967). Differentiation and integration in complex organizations. *Administrative Science Quarterly, 12*, 1-47.
- LeBreton, J. M., Wu, J., & Bing, M. N. (2008). The truth(s) on testing for mediation in the social and organizational sciences. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Received doctrine, verity, and fable in the organizational and social sciences* (pp. 107-141). New York, NY: Routledge.
- MacCallum, R., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*, 19-40.
- MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review, 55*, 95-107.
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Hillsdale, NJ: Erlbaum.
- MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83-104.
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*, 99-128.
- Mathieu, J. E., DeShon, R. P., & Bergh, D. D. (2008). Mediation inferences in organizational research; Then, now, and beyond. *Organizational Research Methods, 11*, 203-223.
- Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12*, 23-44.
- Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin, 113*, 181-190.
- Miller, T. L., Triana, M., Reutzell, C. R., & Certo, S. T. (2007). Mediation in strategic management research: Conceptual beginnings, current application, and future recommendations. In D. Ketchen & D. D. Bergh (Eds.), *Research methodology in strategy and management* (Vol. 4, pp. 295-318). Bingley, UK: Emerald Group.
- Mitchell, T. R., & James, L. R. (2001). Building better theory: Time and the specification of when things happen. *Academy of Management Review, 26*, 530-547.
- Mooney, C. Z., & Duval, R. D. (1993). *Bootstrapping: A nonparametric approach to statistical inference*. Newbury Park, CA: Sage.
- Ndofor, H. A., Sirmon, D. G., & He, X. (2011). Firm resources, competitive actions and performance: Investigating a mediated model with evidence from in-vitro diagnostics industry. *Strategic Management Journal, 32*, 640-657.
- Podsakoff, P. M., & Dalton, D. R. (1987). Research methodology in organizational studies. *Journal of Management, 13*, 419-441.
- Rogers, W. M. (2002). Theoretical and mathematical constraints of interaction regression models. *Organizational Research Methods, 5*, 212-230.
- Schoonhoven, C. B. (1981). Problems with contingency theory: Testing assumptions hidden within the language of contingency theory. *Administrative Science Quarterly, 26*, 349-377.

- Semadeni, M., Withers, M. C., & Certo, T. (2014). The perils of endogeneity and instrumental variables in strategy research: Understanding through simulation. *Strategic Management Journal*, *35*, 1070-1079.
- Shen, W., Kiger, T. B., Davies, S. E., Rasch, R. L., Simon, K. M., & Ones, D. (2011). Samples in applied psychology: Over a decade of research in review. *Journal of Applied Psychology*, *96*, 1055-1064.
- Shook, C. L., Ketchen, D. J., Jr., Hult, G. T. M., & Kacmar, K. M. (2004). An assessment of the use of structural equation modeling in strategic management research. *Strategic Management Journal*, *25*, 397-404.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290-312). San Francisco, CA: Jossey-Bass.
- Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, *89*, 845-851.
- Stine, R. (1989). An introduction to bootstrap methods. *Sociological Methods & Research*, *18*, 243-291.
- Stone-Romero, E. F., & Rosopa, P. J. (2008). The relative validity of inferences about mediation as a function of research design characteristics. *Organizational Research Methods*, *11*, 326-352.
- Stone-Romero, E. F., & Rosopa, P. J. (2011). Experimental tests of mediation: Prospects, problems, and some solutions. *Organizational Research Methods*, *14*, 631-646.
- Thompson, J. D. (1967). *Organizations in action*. New York, NY: McGraw-Hill.
- Vancouver, J. B., & Carlson, B. W. (2015). All things in moderation, including tests of mediation (at least some of the time). *Organizational Research Methods*, *18*, 70-91.
- Wiersema, M. F., & Bowen, H. P. (2009). The use of limited dependent variable techniques in strategy research: Issues and methods. *Strategic Management Journal*, *30*, 679-692.
- Wood, R. E., Goodman, J. S., Beckmann, N., & Cook, A. (2008). Mediation testing in management research: A review and proposals. *Organizational Research Methods*, *11*, 270-295.

Author Biographies

Herman Aguinis is the John F. Mee Chair of Management and the founding and managing director of the Institute for Global Organizational Effectiveness in the Kelley School of Business, Indiana University. His research interests span several human resource management, organizational behavior, and research methods and analysis topics. He has published five books and more than 130 articles in refereed journals. He is a fellow of the Academy of Management, past editor of *Organizational Research Methods*, and received the Academy of Management Research Methods Division Distinguished Career Award.

Jeffrey R. Edwards is the Belk Distinguished Professor of Organizational Behavior at the Kenan-Flagler Business School, University of North Carolina. He is past editor of *Organizational Behavior and Human Decision Processes*, past chair of the Research Methods Division of the Academy of Management, and a fellow of the Academy of Management, the American Psychological Association, and the Society for Industrial and Organizational Psychology. His methodological research addresses difference scores, polynomial regression, moderation and mediation, structural equation modeling, construct validity, and the development and evaluation of theory.

Kyle J. Bradley is a doctoral candidate in organizational behavior and human resource management in the Kelley School of Business, Indiana University. His scholarly interests include research methods, performance management, star performers, and the work-life interface. His work has appeared in *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *Organizational Research Methods*, and *Organizational Dynamics* and has been presented at the meetings of the Academy of Management and Society for Industrial and Organizational Psychology.