Journal of Management
Vol. 44 No. 1, January 2018 131–160
DOI: 10.1177/0149206317690586

Control Variables in Leadership Research: A Qualitative and Quantitative Review

Jeremy B. Bernerth
San Diego State University
Michael S. Cole
Texas Christian University
Erik C. Taylor
Louisiana State University
H. Jack Walker
Auburn University

Statistical control of extraneous (i.e., third) variables is a common analytic tool among leadership researchers. While such a strategy is typically assumed to prove beneficial, it can actually introduce various complications that are underestimated or even ignored. This study investigates and summarizes the current state of control variable usage in leadership research by qualitatively and quantitatively examining the use of statistical control variables in 10 highly regarded management and applied psychology journals. Despite available "best practices," our results indicate that control variable usage in existing leadership studies is rarely grounded in theory but instead frequently relies on outdated misconceptions. Moreover, a meta-analysis of the relationships between popular control variables and leadership constructs finds nearly universal weak effect sizes, suggesting that many studies may not only be losing valuable degrees of freedom but also making inferences based on biased parameter estimates. To address these issues, we put forth a number of recommendations to assist leadership scholars with determining whether potential third variables should be controlled for in their leadership research.

Keywords: *leadership; statistical control; nuisance; covariate; research methodology; survey research*

Acknowledgment: This article was accepted under the editorship of Patrick M. Wright.

Corresponding author: Jeremy B. Bernerth, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-8230, USA.

E-mail: jbernerth@mail.sdsu.edu

Leadership is a critical field of inquiry in the organizational sciences, one that captures the curiosity of a large number of micro-, meso-, and macro-organizational scholars (Gardner, Lowe, Moss, Mahoney, & Cogliser, 2010). As evidence of this, an analysis of submitted manuscripts to one of the premier journals in the management and applied psychology fields found that leadership studies were the second-most popular submission, behind only a catch-all category of "groups/teams" (Morrison, 2010). Whereas authors of leadership studies represent a range of academic disciplines (e.g., education, management, nursing, public administration, psychology, sociology) and have different scholarly interests, they all share a common desire to better understand leadership as an influencing process. Toward this end, leadership researchers frequently need to use nonexperimental research designs to capture the leadership phenomena of interest (Hiller, DeChurch, Murase, & Doty, 2011). Unfortunately, an important limitation of nonexperimental research is that extraneous (i.e., third) variables may exist and produce "distortions in observed relationships" (Spector & Brannick, 2011: 288). For example, a third variable may influence or "contaminate" measures of interest, not by unduly influencing the theoretical constructs, but rather by influencing how individuals respond to survey items (i.e., contaminating assessment). A third variable may also produce a "spurious" relationship between a predictor and criterion such that the two variables correlate as a consequence of sharing a common cause and not because there is a true causal connection between them (i.e., confounded relationship). Finally, there is also a possibility that a third variable relates to a criterion variable but a researcher is interested in studying the unique relationship between a predictor and criterion above a third variable (i.e., incremental validity).

In an attempt to yield more accurate estimates of leadership effects, it is therefore common practice for researchers to use a statistical approach to control for third variable concerns.[1] Often dubbed the "purification principle" (Spector & Brannick, 2011), the notion behind control variable usage is that a researcher can statistically remove any distortions, legitimate or otherwise, associated with extraneous variables, thereby purifying results and exposing "true relationships" (Atinc, Simmering, & Kroll, 2012: 59). There exist, however, a number of well-known complications (at least within the methodological literature) when including control variables in study analyses. For instance, control variables substantively change the meaning of the relationship under investigation such that including a control variable along with a leadership predictor essentially replaces the leadership variable with a new residual predictor (Breaugh, 2006; Edwards, 2008). As this new residual predictor may represent something absent experience, education, or personality (i.e., commonly used control variables), one might reasonably question whether such "fictional" leaders (Meehl, 1970) truly exist. The inclusion of control variables also reduces available degrees of freedom and lowers statistical power, and it may diminish the amount of explainable variance in outcomes attributed to focal predictors (Becker, 2005; Carlson & Wu, 2012). In extreme cases, a control variable may explain so much variance in criteria that a focal predictor may appear unrelated to study outcomes (Breaugh, 2006). Less well known is that the inclusion of a control variable can also increase the likelihood of finding a significant relationship between a predictor and outcomes (i.e., the product of a suppression effect; MacKinnon, Krull, & Lockwood, 2000). As these examples illustrate, ambiguous and even conflicting research findings can legitimately result from the inclusion or exclusion of statistical control variables (Becker, 2005; Breaugh, 2008).

With the above issues in mind, organizational decisions that have significant personal and financial implications are often made as a result of empirical research put forth by leadership scholars, which may inadvertently include methodological deficiencies (see Aguinis & Vandenberg, 2014; Zaccaro & Horn, 2003). Given such realities, one might be surprised to find that there is, to date, no comprehensive review of control variable usage in the context of leadership research. As such, important questions remain unanswered. Is the leadership literature heeding best practices set forth by methodologists and data analysis experts? What specific variables do leadership researchers associate with third variable concerns, and do those variables actually correlate with a study's focal leadership construct? Do different streams of leadership research rely on different justifications when including control variables in statistical tests? The purpose of this study is to answer these and other important questions by connecting the methodological work on control variables with the study of leadership.

Leadership Research and Control Variables: Searching for Answers

As alluded to in the introduction, the metatheory for including statistical controls is best known as the purification principle (for a detailed discussion, see Spector & Brannick, 2011). This meta-explanation includes the term *purification* because there exists an implicit belief among researchers that including statistical controls somehow yields more accurate estimates of predictor-criterion relationships and/or purifies results of alternative explanations. According to Spector and Brannick (2011: 288), this general belief "is so widespread, and is so accepted in practice" that it "qualifies as a methodological urban legend—something accepted without question because researchers and reviewers of their work have seen it used so often that they do not question the validity of the approach." Despite the belief that the purification principle is in line with best practices, methodologists offer detailed mathematical formulae and context-specific examples demonstrating that the use of control variables can produce biased parameter estimates, inferential errors, and a host of other problems (see, e.g., Becker et al., 2016; Breaugh, 2008; Spector & Brannick, 2011). Given the seriousness of these potential complications, multiple pedagogically oriented articles offering recommendations regarding the use of statistical controls exist in the literature.

One fundamental recommendation offered in several of these articles is for researchers to incorporate conceptually meaningful control variables, as opposed to surrogate or proxy variables, which are often used in their place (Becker, 2005; Breaugh, 2008; Spector & Brannick, 2011). If researchers are, for instance, interested in studying leadership emergence, they may wish to control for relationship quality among group members, believing that the affect shared among members might somehow influence who is ultimately nominated as leader. It follows that using department tenure as a surrogate for relationship quality is not only imprecise but may capture other variables unrelated to relationship quality but related to leadership emergence (e.g., company- and job-specific knowledge). A related complication involves the interpretation and replication of a study's results when demographic proxies are used in the data-analytic model. Take the interpretation of a transformational leadership study that statistically controls for managers' age, gender, education, and organizational tenure. Such a study is now investigating an ageless, gender-neutral individual with no education or prior organizational experience, as a consequence of parsing out the variance attributed

to these four factors. Does such an individual exist in reality? Moreover, if a follow-up study does not statistically control for these same four demographic characteristics in its substantive tests, it becomes difficult to compare and generalize results from both studies. Methodologists have indeed suggested that the use of statistical controls make interpretation and comparison across studies more, not less, difficult (e.g., Breaugh, 2008).

To be sure, the methodological literature is quite clear on when and why a control variable should be incorporated into one's statistical analyses (for a review, see Becker et al., 2016). Nevertheless, no existing research within the leadership domain gives explicit attention to the issue of statistical control, nor have prior researchers sought to examine whether leadership scholars heed best practice recommendations. A quick glance at previously published studies suggests that this may represent a significant problem, as scholars appear to use vastly different control variables despite looking at similar relationships. As an example, Bacha (2014) studied the relationship between transformational leadership and follower performance, controlling for leader gender and firm size. Charbonnier-Voirin, Akremi, and Vandenberghe (2010) were also interested in the relationship between transformational leadership and follower performance but instead chose to statistically control for follower demographic factors (i.e., follower age, sex, educational level, and tenure). Still other scholars investigating the relationship between transformational leadership and follower performance forewent control variables altogether (e.g., Choudhary, Akhtar, & Zaheer, 2013; McMurray, Islam, Sarros, & Pirola-Merlo, 2012). As discussed above, these analytic choices have the potential to adversely affect what we know about leadership insomuch as inconsistent practices hinder replicability and fail to guide future research endeavors (Aguinis & Vandenberg, 2014). Thus, we first sought to identify the variables previous leadership research utilized as control variables.

Research Question 1: What variables are statistically controlled for in the study of leadership constructs?

Perhaps just as important as asking "What variables are controlled for?" is asking "Why do researchers believe that these variables should be statistically controlled for in their leadership studies?" To this point, Becker et al. (2016: 158) note, "When in doubt, leave them out." In making this recommendation, Becker and colleagues specifically suggest excluding control variables that lack a clear or defensible purpose, adding that a convincing justification is needed prior to initiating a study. What is a clear and defensible purpose? Existing research does not answer this question explicitly, but several methodologists note that the basis for inclusion *should* lie in theory (Becker, 2005; Bernerth & Aguinis, 2016; Breaugh, 2006). Emphasis is placed on *should* in the previous sentence because generalized reviews across micro- and macrofields indicate that many control variables are covariates of unknown theoretical relevance that act as little more than a "placeholder" (Carlson & Wu, 2012: 418). Conceptually meaningful control variables, however, include constructs embedded within a theory that represent either an established theoretical relationship or a competing theoretical explanation for the anticipated relationship between a leadership predictor and a criterion (Becker et al., 2016). By way of illustration, if theory and prior empirical evidence suggests that a particular leadership variable (e.g., initiating structure) relates to the criterion and if a researcher is interested in determining whether another leadership variable (e.g.,

consideration) has a unique relationship with that same criterion, including the first variable as a statistical control has a clear purpose and is therefore defensible.

The extent to which theory guides control variable decisions within leadership research is unknown, but personal experience suggests the use of many other, nontheoretical justifications for including control variables. An inspection of recent articles published in management and applied psychology journals reveals several such atheoretical justifications for the use of statistical controls, including the pursuit of a more stringent test of one's proposed hypotheses and striving to be consistent with previous research. As such, we next sought to identify the reasons why scholars include control variables in leadership research.

Research Question 2: (a) In the leadership literature, to what extent is the inclusion (or exclusion) of statistical control variables grounded in theory? (b) If not grounded in theory, why are statistical control variables included in leadership research?

Although careful selection and sound justification are important, methodologists recommend several other best practices regarding control variable treatment. Recently, Becker et al. (2016) integrated and condensed these recommendations into four broad categories, three of which take us beyond control variable selection. When one or more control variables are incorporated into a study's design, analyses, and results, the authors urge researchers to *include control variables in their formal hypotheses*. Assuming that there is a theoretical justification for control variable use, they assert that including control variables into the corresponding hypothesis helps to ensure the appropriate test of theory. Consider the frequent scenario in which a study's formal hypotheses are stated in terms of bivariate relationships but then tested with control variables in the analyses. According to Becker et al., this situation results in an invalid test of the hypotheses because proposed relationships (as formally stated) have yet to be assessed.

Another broad topic relates to *reporting and interpreting results*. Specifically, Becker et al. (2016) urge the reporting of standard descriptive statistics and correlations for control variables. Doing so provides an opportunity to better understand the psychometric properties of control variables, which in turn enhances the potential for replication. Additionally, by including control variables in the intercorrelation matrix, one can calculate the amount of shared variance between a study's focal variables and control variables. This is a potentially important piece of information because one can subsequently gauge the amount of residual variance that remains in a leadership predictor after the shared or common variance associated with the control variables has been parsed. To the extent that knowledge of residual variance may aid in the interpretation of more sophisticated data analyses (e.g., why hierarchical regression analyses report an insignificant coefficient for a predictor sharing a significant bivariate relationship with a criterion), it serves science and the field of leadership to report all descriptive information for every control variable included or excluded in a study.

One other area relates to *measuring and analyzing control variables*. Previous work demonstrates that when control variables are measured with error, parameter estimates of substantive relationships can become upwardly or downwardly biased (Edwards, 2008). Becker et al. (2016) emphasize, therefore, that reliability information (i.e., coefficient alpha) must be reported when control variables are perceptual in nature, as this practice ensures that measurement error is within reason and/or may be corrected. Along with providing reliability information, researchers are advised to analyze and report study results with and without

control variables; this helps illustrate the impact of control variables on the relationships between focal predictors and criteria.

In sum, the quality of published research is enhanced when (a) control variables are incorporated into a study's hypotheses, (b) standard descriptive statistics and correlations are reported for control variables, (c) control variables are subjected to the same standards of reliability as other variables, and (d) researchers openly compare results with and without control variables (Becker et al., 2016). Scholars well read in the social sciences likely have a sense of the commonality (or rarity) of such practices, yet we are unaware of any systematic effort within the leadership area to assess the state of the field. Without such knowledge, it is difficult to ascertain if leadership studies are properly using control variables. We therefore sought to assess the quality of leadership research vis-à-vis the frequency with which these essential best practices are adopted in prior published work.

Research Question 3: Do leadership scholars (a) include control variables in study hypotheses, (b) report all descriptive statistics and correlations for control variables, (c) assess the reliability of control variables, and (d) compare results with and without control variables?

One final focal concern of methodologists is the relationship between control variables and a study's independent and dependent variables. These relationships are important for a number of reasons. For example, due in part to the shared variance between a control variable and an independent variable, incorporating statistical controls into one's analyses can artificially increase the magnitude of a regression coefficient (for a detailed discussion of confounding and suppression effects, see MacKinnon et al., 2000), leading to the imaginable scenario in which one concludes that the leadership predictor relates to the dependent variable when in fact it does not (i.e., Type I error via suppression). Another potential complication arises when control variables are incorporated for the sake of "purification" and explain so much variance in the dependent variable that the leadership predictor is unable to explain a significant amount of variance even when a "true" relationship does in fact exist (i.e., Type II error). Becker and colleagues (Becker, 2005; Becker et al., 2016) also discuss the implications of entering into analyses "impotent" controls (i.e., those control variables that have no relationship with the dependent variable), including a reduction in statistical power as a result of losing degrees of freedom.

It follows that having a better understanding of the relationships between common control variables and leadership constructs serves several purposes. When leadership variables act as the predictor, which is frequently the case, established relationships serve as an indication of how much variance in a leadership construct remains prior to the assessment of the predictor-criterion relationship. Knowing that, on average, only 65% of an original leadership construct remains after removing the influence of a popular control variable subsequently gives context to regression coefficients and explainable variance. As such, establishing relationships between commonly used control variables and leadership constructs offers possible preventative medicine against the inappropriate generalization of results based on residual predictors (Edwards, 2008), the waste of degrees of freedom (Becker, 2005), the provision of uninterpretable results (Spector & Brannick, 2011), and other complications associated with the use of control variables (Becker et al., 2016). For such reasons, our final research question explores these relationships.

Research Question 4: In the leadership literature, to what extent are statistical control variables correlated with focal leadership constructs?

Method

Literature Review

According to Gardner et al. (2010) and Dinh et al. (2014), among others, the full-range leadership model (e.g., Avolio & Bass, 1990; Bass, 1985) and leader-member exchange (LMX; e.g., Graen & Scandura, 1987) are the dominant perspectives from which to understand leadership effects. Our initial database search (i.e., ABInform; EBSCOhost) therefore focused on the following keywords: *transformational leadership* (and *TFL*), *idealized influence*, *charisma*, *inspirational motivation*, *intellectual stimulation*, *individualized consideration*, *transactional leadership*, *contingent-reward*, *active management-by-exception*, *passive management-by-exception*, *laissez-faire leadership*, and *leader-member exchange* (and *LMX*). Moreover, to broaden the scope of our study, we examined narrative reviews (e.g., Day & Antonakis, 2011) and conducted a secondary database search to identify other routinely studied leadership constructs. These supplementary searches, an approach that is consistent with prior reviews of the leadership literature (e.g., Dinh et al., 2014), suggested the inclusion of three additional leadership constructs: *authentic leadership* (e.g., George, 2003), *ethical leadership* (e.g., M. E. Brown & Treviño, 2006), and *shared leadership* (e.g., Carson, Tesluk, & Marrone, 2007; Wang, Waldman, & Zhang, 2014).

For readers unfamiliar with some (or all) of these topics, we note that the full-range leadership model—which represents a merger of traditional and modern paradigms of leadership—is composed of transformational, transactional, and laissez-faire leadership behaviors. Transformational leadership focuses largely on the behaviors and attributes of leaders that inspire followers to set aside their self-interests and instead focus on the greater good. Facets of transformational leadership include attributed (i.e., charisma) and behavioral idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration. In contrast to transformational leadership, transactional leadership focuses on an exchange process used to motivate and ensure follower compliance with organizational rules and norms (Yukl, 1999). Specific facets include contingent-reward behaviors, which clarifies role requirements and the rewards associated with successful task compliance, and active management-by-exception, which represents close monitoring and adjustment of followers' behaviors (cf. Bono & Judge, 2004). Laissez-faire leadership, which we grouped with passive management-by-exception (cf. DeRue, Nahrgang, Wellman, & Humphrey, 2011), largely represents an absence of leadership.[2] Relational leadership models, which are frequently studied under the umbrella of LMX, focus on the social exchanges that occur between leaders and followers (Bernerth, Armenakis, Feild, Giles, & Walker, 2007).

What we describe as "emerging" leadership domains represent leadership styles that have yet to receive as much attention as the full-range or relational models of leadership but are nonetheless increasingly studied. Authentic leadership focuses on the extent to which leaders are aware of both their abilities and short-comings and lead with a strong sense of their values and principles (George, 2003), while ethical leadership emphasizes a moral component of leadership through the communication of and compliance with ethical values and standards (M. E. Brown & Treviño, 2006). Shared leadership, a final emerging area, focuses on

the sharedness of leadership and distribution of influence throughout a team (Wang et al., 2014).

On the basis of previous examinations of journal quality (Zickar & Highhouse, 2001) with the most recent journal impact data published by the ISI Web of Knowledge, we searched 10 journals for published research on these topics. Specifically, we included *Academy of Management Journal*, *Administrative Science Quarterly*, *Group & Organization Management*, *Journal of Applied Psychology*, *Journal of Business and Psychology*, *Journal of Organizational Behavior*, *Journal of Management*, *The Leadership Quarterly*, *Organizational Behavior and Human Decision Processes*, and *Personnel Psychology*. Within these journals, we sought to identify leadership studies published during 2003 to 2014. We selected this time frame because it allows for several years of published research that predates influential methodological articles on statistical control (e.g., Becker, 2005; Breaugh, 2006; Edwards, 2008) and captures the most recent decade of published studies on leadership effects. Given our study's purposes, we excluded meta-analytic studies and laboratory experiments. We also excluded qualitative examinations of leaders' speeches, press releases, and media coverage.

Process for Identifying Studies

We used a four-step process to identify primary studies for inclusion in our analyses. This process began with an online keyword search of abstracts. We then read the article's abstract to ensure the appropriateness of the study; that is, we eliminated those studies referencing qualitative methodologies and meta-analyses. Next, we examined the article's Method section to determine if statistical controls were incorporated into the study's design. If we did not identify any statistical controls in an article's Method section, we then performed a within-article keyword search of the terms "control" and "covariate." We also examined "results" tables to identify variables not described elsewhere in the article but were clearly used as controls. Combined, these efforts resulted in the identification of 290 empirical articles that (a) examined at least one of the focal leadership styles and (b) incorporated at least one variable identified as a "control" or "covariate" in its analyses.

Categorizing Leadership Domains

In categorizing and coding transformational leadership, we included studies that reported using a composite (i.e., global) measure of transformational leadership as well as those studies that used one or more of transformational leadership's subdimensions (cf. Judge & Piccolo, 2004). Consistent with Bono and Judge (2004), our transactional leadership variable comprises studies that reported using either a composite measure of transactional leadership or one of its two dimensions (i.e., contingent-reward and active management-by-exception); laissez-faire leadership included either or both laissez-faire and passive management-by-exception measures (cf. Bono, Hooper, & Yoon, 2012; Bono & Judge, 2004; DeRue et al., 2011).

Coding Issues, Decisions, and Procedures

To answer Research Question 1, we recorded each control variable used in the study of the seven leadership domains at the most inclusive level, resulting in the identification of control

variables both specific (e.g., leader age, gender) and general (e.g., job function, performance). To assess Research Questions 2a and 2b, we needed to code the various justifications given for using control variables. Following best practice recommendations (Duriau, Reger, & Pfaffer, 2007), two authors independently (a) reviewed the justifications given for the use of roughly 1,400 statistical control variables and (b) generated a list of common themes believed to encompass the justifications. Through discussion, the two authors subsequently agreed that 12 justification codes and themes adequately captured the underlying rationales provided in the primary studies. We initially coded for (1) the presence (absence) of any type of justification (labeled *explanation* from this point forward) and, (2) when one existed, whether the justification was theoretical (nontheoretical) in nature (labeled *theoretical*). To qualify as theoretical, the provided justification had to explain the what, the how, and the why regarding the relationship between the control variable and the focal variable (see Bacharach, 1989; Sutton & Staw, 1995). The additional coding themes included (3) the presence (absence) of a citation for the justification/description of control variables (labeled *citation*), (4) a potential relationship between the control variable and a study's focal variable (labeled *potential*), (5) previously found empirical relationships between the control variable and a study's focal variable (labeled *previous*), (6) prior research including such variables as controls (labeled *imitation*), (7) a desire to eliminate alternative relationships (sometimes but not always referencing spurious or confounding ones; labeled *alternative*), (8) the explicit reference to method bias (labeled *method bias*), (9) a belief that such controls enhanced the robustness of their results (sometimes but not always referencing conservative or stringent tests; labeled *robustness*), (10) an effort to offer evidence of incremental or discriminant validity (labeled *validity*), (11) an analysis of a study's primary data (e.g., potential controls were found to significantly relate to focal variables and, as a result, included; labeled *analysis*), and last, (12) whether researchers utilized a multistep process for deciding whether to include control variables (e.g., combining previous empirical findings with the analysis of one's own data characteristics; labeled *process*). Examples of each of these 12 categories appear in Table 1.

Research Questions 3 and 4 also required us to make additional coding decisions. Specifically, Becker et al. (2016) offer several reporting recommendations when one or more control variables are incorporated into a study's research design and analyses. As such, we recorded whether a study included control variables in its hypotheses. We coded whether each primary study reported standard descriptive statistics (means, standard deviations, and defining information, such as 1 = males, 2 = females) and the correlations between control variables and focal variables. If a study included perceptual control variables (e.g., positive affect), we also recorded whether reliability information was provided for that variable. We likewise coded whether a study reported running analyses with and without control variables. Finally, when correlations between control variables and focal leadership variables were reported, we recorded (a) statistical significance (i.e., 1 = significant, 0 = not significant, based on whether authors indicated the relationship was significant) and (b) the actual effect size reported.

Interrater Agreement

In regard to the 12 justification categories (see Table 1), two authors independently coded each control variable included in the 290 identified studies. Interrater agreement was 87% or

Table 1
Control Variable Justification Categories, Coding, and Examples

Justification	Coding	Examples
Justification	0 = no justification given 1 = some justification given	Employees' age and organizational tenure were included as control variables. Coded as 0. Because previous research has shown demographic variables as well as demographic similarity may influence the leader-subordinate relationship and subsequent performance (McColl-Kennedy & Anderson, 2005; Turban & Jones, 1988), we controlled for subordinate and supervisor gender and age, as well as differences in gender and age. Coded as 1. *Source*: Bernerth et al. (2007).
Theoretical	0 = justification did not explain how/why controls relate to a study focal variable 1 = justification of control explains how/why control relates to a study focal variable	. . . Since larger spans of control can diminish a leader's ability to influence, it is likely that a wide span of control would lead to fewer resources allocated across employees. Similar arguments have been made with respect to employee performance ratings (e.g., Judge & Ferris, 1993). Much as professors who teach large sections have less time to devote to individual student needs, leaders with wider spans of control also have less time to engage employees with transformational leadership behavior such as providing individualized support or intellectual stimulation. Coded as 1. *Source*: Rubin et al. (2005). . . . we controlled for dispositional general affectivity of the nurses in the analyses. Dispositional general affectivity reflects pervasive individual differences that systematically bias people to express positive or negative evaluations of the various aspects of their environment (Cropanzano, James, & Konovsky, 1993; Watson, Clark, & Tellegen, 1988). Our independent variable (i.e., LMX) and the dependent variables (e.g., POS) were nurses' evaluations of various facets of their organization and work relationships. Hence, controlling for general affectivity of the nurses allowed us to factor out common method variance that can be ascribed to such affectivity (i.e., their disposition to positively or negatively respond to survey items) in our analyses. Coded as 1. *Source*: Tangirala et al. (2007). . . . In addition, leaders supervising smaller groups of individuals have more time and energy to devote to relationship building than do leaders supervising larger groups (Schriesheim, Castro, & Yammarino, 2000; Schyns, Paul, Mohr, & Blank, 2005). Given the differences in the number of subordinates reported by the leaders in our sample, we also controlled for the leader's span of supervision with number of employees. Coded as 1. *Source*: Moss et al. (2009). One of the key assumptions of both models proposed above is that charisma is an antecedent of performance. Cumulative evidence supports this formulation (e.g., Judge & Piccolo, 2004). However, there is the possibility that subordinates attribute charisma to a leader because the leader is a proven high performer (Agle, Nagarajan, Sonnenfeld, & Srinivasan, 2006; Keller, 1992; Meindl, 1995). To help control for this possibility, we assessed the subordinates' perceptions of the prior effectiveness of team leaders using . . . Coded as 1. *Source*: Balkundi et al. (2011).
Citation	0 = justification does not offer a citation 1 = justification includes a citation	To control for the possibility that third variables might lead to spurious relationships, gender, age, and organizational tenure were entered as covariates in the analyses. Coded as 0. We controlled subordinate sex, age, tenure, and education, which could covary with independent and dependent variables (Zellars et al., 2002). Coded as 1. *Source*: Xu, Huang, Lam, and Miao (2012).
Potential	0 = does not justify control by explicitly referencing potential relationships 1 = justifies control by explicitly referencing potential relationships	We controlled for employee organizational tenure because organizational experience may relate to understanding norms for performance and therefore may be related to individual and group performance. Moreover, employees with greater organizational tenure may find it easier to form high quality relationships with their leaders. Coded as 1. *Source*: Liden, Erdogan, Wayne, and Sparrowe (2006). … older participants might have more experience dealing with bosses, which could influence their perceptions of leadership. Coded as 1. *Source*: Pastor, Mayo, and Shamir (2007).

(continued)

Table 1 (continued)

Justification	Coding	Examples
Previous	0 = does not reference previous empirical findings to justify control 1 = references previous empirical findings to justify control	Previous research indicates these variables relate to our outcome variables (e.g., supporting citations) Coded as 1. We included four control variables that prior research has linked with perceptions of leadership. Coded as 1.
Imitation	0 = does not justify control by referencing others who used it as a control 1 = justifies control by referencing others who included it as a control	Consistent with prior research, we controlled for . . . in all our analyses to reduce alternative explanations. Coded as 1. We included three control variables suggested by prior research. First, . . . Coded as 1.
Alternative	0 = does not justify control by explicitly referencing the elimination of alternative relationships 1 = justifies control by explicitly referencing the elimination of alternative relationships	Individuals' educational level, age, sex, and number of previous jobs were included as control variables in our analyses to rule out potential alternative explanations for our findings. Research has suggested that. . . . It is possible that these individual differences also shape individuals' perceived investments and costs in developing an exchange relationship, influencing their readiness to make themselves vulnerable and thus, trust their organization. Coded as 1. *Source*: Dulac, Coyle-Shapiro, Henderson, and Wayne (2008). We controlled for age, sex, education, and organization tenure to avoid potential confounding effects on our dependent variables. Coded as 1.
Method bias	0 = does not reference method bias 1 = references method bias explicitly	Additionally, because study variables were self-reported, we measured socially desirable responding with a short form of the Marlow–Crowne Social Desirability Scale. Coded as 1. Controlling for these variables minimized the effects of common method biases on our results (cf. Podsakoff et al., 2003). Coded as 1.
Robustness	0 = does not reference robust or conservative tests of hypotheses 1 = explicitly references robust or conservative test of hypotheses	Four demographic variables were used as controls in the supplementary analyses to examine the robustness of our findings. Coded as 1. To provide more accurate estimates of our hypothesized relationships, we controlled for . . . Coded as 1.
Validity	0 = does not reference incremental or discriminant validity 1 = justifies control with an incremental or discriminant validation explanation	We included a statistical control for transactional leadership in all analyses predicting ratings of charisma and a control for ratings of charisma in our analyses predicting transactional leadership. Even though they are conceptually different, charismatic and transactional leadership are highly correlated in the leadership literature, with most studies reporting raw correlations between transformational and transactional leadership in the .70s. For instance, in a recent meta-analysis, Judge and Piccolo (2004) reported a meta-analytic estimated correlation between transformational and transactional leadership of $\rho = .80$. Besides, we have argued here that arousal affects ratings of charisma but not ratings of transactional leadership in general. If we show that ratings of charisma are influenced by arousal after controlling for transactional leadership, we have strong inferential support for our argument. That is, by controlling for ratings of transactional leadership and the variance they share with ratings of charisma, we have removed from the analysis the possible effects of arousal on perceptions of leadership in general. In a similar vein, we included a control for ratings of charisma when predicting ratings of transactional leadership. Coded as 1. *Source*: Pastor et al. (2007). We also measured interactional justice (1 = strongly disagree; 5 = strongly agree; $\alpha = .85$; Niehoff & Moorman, 1993) to examine whether LMX could explain additional variance above and beyond interactional justice, an established mediator of the link between abusive supervision and work behaviors. Coded as 1. *Source*: Xu et al. (2012).

(continued)

Table 1 (continued)

Justification	Coding	Examples
Analysis	0 = justification did not describe an analysis of study data 1 = justified control decision with analysis of study data	As shown in this table, few demographic variables were related to endogenous variables in the model. However, because subordinate age was positively related to LMX ratings, we controlled for the effects of age on LMX ratings in our model tests. No other demographic variables were included in the model. Coded as 1. *Source*: Dahling, Chau, and O'Malley (2012). Because participants in this study had relatively long average tenure with the organization, we analyzed data with and without controlling for organizational tenure, and the results were almost identical. Thus, we do not include organizational tenure as a control variable. Coded as 1. *Source*: Sekiguchi, Burton, and Sablynski (2008).
Process	0 = does not describe (at a minimum) a two-step process for control inclusion/exclusion 1 = describes (at a minimum) a two-step process for control inclusion/exclusion	Consistent with prior research on LMX (e.g., Bauer & Green 1996; Bernerth et al. 2008; Green et al. 1996; Liden et al. 1993), we also tested several demographic variables as potential control variables in our analysis, including supervisor and subordinate gender (0 = male; 1 = female), age, and education levels (0 = junior high school diploma to 5 = graduate degree), as well as supervisor–subordinate differences on these variables. Correlation analyses indicated that these variables were not significantly associated with EI or LMX; therefore, to maintain model parsimony, these demographic variables were not included in subsequent analyses. Coded as 1. *Source*: Sears and Holmvall (2010). Our potential control variables included gender, age, racioethnicity (Caucasian versus minority), previous job experience (full-time and part-time), full-time (versus part-time) employment status, permanent (versus probationary) employment status, and inbound (versus outbound) employment status. While gender, age, and racioethnicity are common control variables, work experience and employment status also may influence adjustment. More work experience may negate the influence of socialization sources (e.g., organization, supervisory; Saks et al., 2007) whereas employment status may precondition organizational and occupational attitudes. Nevertheless, Becker (2005, p. 285) warns of the use of ''impotent control variables'' that may unnecessarily reduce power. Thus, we only include those variables that predict significant variance in the relevant endogenous variable (see Kline, 2005). As a result, we use (1) full-time employment status for the prediction of leader–member exchange; (2) age for the prediction of leader–member exchange; and (3) inbound call-center status for the prediction of PO fit and job satisfaction. Coded as 1. *Source*: Sluss and Thompson (2012).

Note: Examples that give a source represent direct quotations. Examples without source information represent generic statements commonly seen in published research. We do not give specific examples in some categories to avoid drawing attention to individual work. We also note some of the direct quotations represent only part of the justification offered in the original work; we quote only part of the justification to demonstrate a specific coding feature. For the full justification, interested readers may see the citation information listed in the references.

better for each unique category, with an overall average agreement of 96%. Where differences existed, we addressed and resolved all discrepancies through discussion. For the more descriptive-type categories examined in relation to Research Question 3, coding was completed by one author, with a random check of 30 studies by a second author. No discrepancies were found in this process.

Data Analysis and Reporting

We begin by presenting a qualitative analysis that describes the number of unique studies including a specific control and the frequency of that particular control variable for given leadership areas (e.g., the percentage of LMX studies using statistical controls that controlled for follower gender). In addition, we provide the percentage of instances that the control variable was correlated with the focal leadership variable and meta-analytic effect sizes associated with the control variables used in the study of LMX, transformational leadership, and transactional leadership.[3] We based this analysis on Hunter and Schmidt's (2004) meta-analytic approach in which sampling error and unreliability are corrected. When a statistical control variable was objective in nature (e.g., gender, age, tenure), we assumed that it was reliably reported; if a primary study did not report reliability estimates, we imputed the average reliability associated with the specific leadership construct. In a few instances when a primary study reported facet-level correlations and not construct-level correlations, we averaged facets and reliability estimates prior to meta-analytic procedures (cf. Bono & Judge, 2004). Finally, if two publications used the same data, we included only one in our meta-analyses and in the counts associated with our research questions.

Results

What Variables Are Statistically Controlled for in the Study of Leadership?

Regarding Research Question 1, Table 2 summarizes the number and percentage of unique studies including a specific statistical control. In terms of the raw number of studies using a specific control variable, results indicate similarities and differences across the various forms of leadership. In terms of similarities, personal demographics (e.g., age and gender) and tenure-related controls appeared in more studies overall and within each focal leadership domain (weighted average inclusion rate = 57%) than any of the other control variables. Within each category, the breakdown of specific aspects included such factors as follower organizational tenure and job tenure, leader organizational tenure and job tenure, follower gender and age, leader gender and age, gender differences, and age differences, among others. Education-related controls (weighted average inclusion rate of 27%) and team/group size controls (weighted average inclusion rate of 21%), along with race-related controls (weighted average inclusion rate of 10%) emerged as the next most frequently occurring categories across leadership areas. Other factors appearing in nearly 20 unique leadership studies across the seven leadership domains included organizational size, organizational level, and job function.

While the general pattern of statistically controlled variables (across leadership domains) mostly converged around demographic variables, Table 2 reveals some differences: 19% of transactional and 30% of laissez-faire leadership studies statistically controlled for

Table 2
Statistical Control Variables Used in the Study of Leadership

	LMX		Transformational		Transactional		Laissez-Faire		Authentic		Ethical		Shared	
Control	*n*	Inclusion Rate[a]	*n*	Inclusion Rate	*n*	Inclusion Rate	*n*	Inclusion Rate	*n*	Inclusion Rate	*n*	Inclusion Rate	*n*	Inclusion Rate
Gender	88	72%	76	52%	15	48%	4	40%	3	33%	7	50%	4	57%
Follower	74	61%	54	37%	12	39%	3	30%	3	33%	6	43%	2	29%
Leader	16	13%	27	18%	5	16%	1	10%	—	—	2	14%	—	—
Gender differences	21	17%	10	7%	—	—	—	—	—	—	—	—	2	29%
Other gender-related controls	—	—	3	2%	—	—	—	—	—	—	—	—	—	—
Tenure	97	80%	80	55%	14	45%	3	30%	4	44%	7	50%	—	—
Follower job	15	12%	7	5%	1	3%	—	—	—	—	3	21%	—	—
Follower organizational	57	47%	37	25%	5	16%	1	10%	3	33%	4	29%	—	—
Team	8	7%	13	9%	—	—	—	—	1	11%	1	7%	—	—
Tenure under leader	48	39%	29	20%	8	26%	1	10%	1	11%	2	14%	—	—
Leader's job tenure	1	1%	14	10%	6	19%	1	10%	—	—	1	7%	—	—
Leader's organizational tenure	10	8%	9	6%	—	—	—	—	—	—	1	7%	—	—
Other tenure-related controls	7	6%	10	7%	—	—	—	—	—	—	1	7%	—	—
Age	82	67%	66	45%	14	45%	2	20%	7	78%	5	36%	5	71%
Follower	71	58%	47	32%	11	35%	2	20%	5	56%	5	36%	1	14%
Leader	14	11%	20	14%	4	13%	—	—	—	—	1	7%	—	—
Age differences	18	15%	6	4%	—	—	—	—	—	—	—	—	1	14%
Organizational	—	—	6	4%	2	6%	—	—	1	11%	—	—	2	29%
Other age-related controls	—	—	4	3%	—	—	—	—	1	11%	—	—	1	14%
Education	40	33%	36	25%	3	10%	1	10%	4	44%	5	36%	1	14%
Follower	37	30%	28	19%	3	10%	1	10%	4	44%	4	29%	1	14%
Leader	11	9%	9	6%	—	—	—	—	—	—	1	7%	—	—
Education differences	7	6%	5	3%	—	—	—	—	—	—	—	—	—	—
Race	15	12%	15	10%	1	3%	2	20%	1	11%	—	—	1	14%
Follower	7	6%	14	10%	1	3%	2	20%	1	11%	—	—	—	—
Leader	—	—	3	2%	—	—	—	—	—	—	—	—	—	—
Racial differences	9	7%	1	1%	—	—	—	—	—	—	—	—	1	14%

(continued)

Table 2 (continued)

Control	LMX		Transformational		Transactional		Laissez-Faire		Authentic		Ethical		Shared	
	n	Inclusion Rate[a]	*n*	Inclusion Rate	*n*	Inclusion Rate	*n*	Inclusion Rate	*n*	Inclusion Rate	*n*	Inclusion Rate	*n*	Inclusion Rate
Organization size	4	3%	21	14%	5	16%	2	20%	1	11%	1	7%	2	29%
Team/group size	27	22%	32	22%	3	10%	—	—	2	22%	2	14%	5	71%
Personality	7	6%	16	11%	1	3%	1	10%	2	22%	—	—	2	29%
Follower Big Five traits	1	1%	3	2%	—	—	—	—	—	—	—	—	—	—
Follower negative affect	5	4%	3	2%	1	3%	—	—	1	11%	—	—	1	14%
Follower positive affect	5	4%	4	3%	—	—	—	—	—	—	—	—	—	—
Other personality-related controls	3	2%	10	7%	—	—	1	10%	1	11%	—	—	1	14%
Follower work experience	8	7%	7	5%	1	3%	—	—	—	—	—	—	—	—
Follower work status	5	4%	7	5%	1	3%	—	—	—	—	2	14%	—	—
Follower workload/hours	6	5%	2	1%	1	3%	—	—	—	—	1	7%	—	—
Follower organization level	13	11%	13	9%	2	6%	1	10%	—	—	—	—	—	—
Family-related controls[b]	5	4%	2	1%	1	3%	1	10%	—	—	—	—	—	—
Follower job function	8	7%	13	9%	—	—	1	10%	—	—	—	—	—	—
Leader span of control	2	2%	5	3%	2	6%	—	—	—	—	2	14%	—	—
Social desirability	1	1%	5	3%	—	—	—	—	—	—	1	7%	—	—
Performance-related controls	2	2%	4	3%		—	—	—	—	—	—	—	—	—
TFL	—	—	—	—	6	19%	3	30%	1	11%	2	14%	—	—
Transactional	—	—	9	6%	—	—	3	30%	—	—	—	—	—	—
Controls relating to sample[c]	18		42		5		1		1		2		—	
Other[d]	31		45		10		2		3		3		4	
Overall total number of studies	122		147		31		10		9		14		7	

Note: Overall number of unique studies, $n = 290$. The number of studies listed under each leadership category are not exclusive. That is, several studies using statistical controls investigated more than one focal topic simultaneously (e.g., TFL and transactional leadership). In such cases, the control(s) was counted and listed under both topics in Table 1. LFX = leader-member exchange; TFL = transformational leadership.

[a]The percentage listed in this column represents the frequency of which a study using control variables included this specific factor (e.g., 72% of LMX studies using statistical controls included some type of gender variable).

[b]Controls included in this composite measure included various follower family-related factors, such as marital status, the number of children living at home, and so on.

[c]The numbers shown in this row represent studies employing controls related to data/sample quality. Where possible, certain factors were collapsed into a single category. This decision was made because researchers often included context-specific control variables that represented similar factors within and between studies (e.g., dummy codes for location, organization).

[d]The numbers shown in this row represent the number of additional control variables not included in any of the other categories listed in the table.

transformational leadership, yet no LMX study included transformational leadership as a control. Moreover, 6% of transformational leadership studies included transactional leadership as a control, while no LMX, authentic, ethical, or shared leadership study included such a control. Another noticeable difference across leadership areas is the emphasis placed on controlling for demographic *differences*. As mentioned above, gender, age, race, and education controls were included frequently across most leadership domains, yet LMX research places a greater emphasis on demographic differences between followers and leaders than other leadership areas. We also found some variation within various tenure- and age-related controls. LMX studies more frequently include follower organizational tenure (47% of studies) and dyadic tenure (i.e., tenure under leader; 39% of studies), while transactional leadership studies include leader's job tenure (19% of studies) more frequently than do other leadership streams. Moreover, LMX (58%) and authentic leadership studies (56%) include follower age more frequently as compared with the other areas.

Why Are Control Variables Included in Leadership Research?

Research Questions 2a and 2b looked to identify the extent to which the inclusion of statistical control variables is based on theory and, if not based on theory, what other reasons explain why leadership scholars control for a specific variable. As shown in Table 3, we provide a content justification analysis across the seven leadership domains and leadership research overall. In relation to Research Question 2a (i.e., to what extent is the inclusion or exclusion of statistical control variables grounded in theory), results indicate that control variable usage is rarely based on theoretical reasons. In fact, despite repeated recommendations that researchers should provide a strong theoretical basis for including each control variable (see Becker et al., 2016, and references therein), only 2% of leadership studies across the seven leadership domains provided a theoretically grounded justification for the inclusion of a specific control variable, yet 73% of the identified studies provided some type of explanation. Upon further inspection, it appears that the most common justification for including covariates is that the authors believed that control variables could relate to a study's focal variable (*potential*, 31%) and/or that prior research has reported a substantive relationship between the control variable and a study's focal variable (*previous*, 22%).

As an aside, we noted some observable differences between the various leadership streams and their justification usage. For example, Table 3 suggests that, compared with the combined average, researchers studying shared leadership (83%), transformational leadership (80%), and laissez-faire leadership (78%) do a relatively better job of offering at least some type of justification. In contrast, ethical leadership (48%), transactional leadership (61%), and authentic leadership (63%) studies were all found to less frequently justify their use of control variables. Other differences appear in the citations offered, as transformational leadership studies provide citations for 56% of control variables, whereas LMX and transactional leadership studies offer citations for only 48% and 41% of control variables, respectively.[4]

What Are the Reporting Habits of Leadership Scholars When Using Control Variables?

Research Questions 3a-3d sought to further examine how well leadership studies adhere to methodologists' advice regarding the appropriate use of statistical controls. Specifically,

Table 3
Control Variable Justifications Across Leadership Domains

Focal Variable	Explanation	Theoretical Explanation	Citation	Potential	Previous	Imitation	Alternative	Method Bias	Robustness	Validity	Analysis	Process	Controls,[a] *n*
Leader member exchange	70%	3%	48%	33%	18%	7%	17%	2%	2%	4%	14%	7%	577
Transformational leadership	80%	2%	56%	29%	27%	13%	17%	2%	3%	2%	14%	5%	585
Transactional leadership	61%	3%	41%	29%	17%	10%	8%	1%	0%	7%	10%	3%	93
Laissez-faire	78%	4%	52%	26%	7%	22%	19%	7%	19%	22%	7%	0%	27
Authentic leadership	63%	0%	40%	30%	20%	3%	13%	0%	0%	0%	0%	0%	30
Ethical leadership	48%	2%	42%	15%	27%	2%	6%	0%	0%	4%	6%	2%	48
Shared leadership	83%	8%	67%	63%	17%	8%	17%	0%	0%	0%	8%	4%	24
Combined leadership studies	73%	2%	51%	31%	22%	10%	16%	2%	3%	4%	13%	6%	1,384

Note: The values shown in this table represent the percentage of studies including particular types of explanations in their control variable justification. Total percentage within each row may not add to 100%, because some studies used multiple justifications that were not mutually exclusive. Explanation = the presence of any type of justification. Theoretical explanation = authors explain what, why, and how they expect a relationship between control variables and focal variables. Citation = the presence of a citation for justification/description of control variables. Potential = justify inclusion by describing a potential relationship between the control variable and a study's focal variable. Previous = justify inclusion by describing a previously found empirical relationship between the control variable and a study's focal variable. Imitation = study cites previous research including such variables as controls. Alternative = described a desire to eliminate alternative relationships (sometimes but not always referencing spurious or confounding ones). Method bias = explicitly references method bias. Robustness = authors felt that including such controls enhanced the robustness of their results (sometimes but not always referencing conservative or stringent tests). Validity = justify inclusion with a reference to incremental or discriminant validity. Analysis = study analyzes the relationship between potential controls and focal variables prior to inclusion/exclusion decisions. Process = researchers utilized a multistep process for deciding whether or not to include control variables (e.g., combining previous empirical findings with the analysis of one's own data characteristics).

[a]This column reflects the total number of control variable instances within each leadership domain. Using this information, one can better compare and contrast the frequency of justification use across leadership constructs.

Table 4
Control Variable Reporting Practices in Published Leadership Research

Focal Variable	Included in Hypotheses[a]	Reported Descriptive Statistics[b]	Reported Reliability[c]	Ran With and Without[d]
Leader member exchange	0%	68%	95%	3%
Transformational leadership	1%	69%	90%	8%
Transactional leadership	0%	65%	87%	6%
Laissez-faire	0%	89%	93%	7%
Authentic leadership	0 %	100%	100%	0%
Ethical leadership	0%	79%	100%	0%
Shared leadership	0%	67%	67%	4%
Combined leadership studies	0%	70%	92%	5%

Note: The values in this table represents the percentage of control variables that have a particular feature.
[a]Control variable included in study hypotheses. The overall percentage shown in this column is zero despite finding a select few cases where traditional controls were included in study hypotheses. This is a result of rounding. Not shown in this column nor table are those cases in which authors either (a) controlled for a focal variable in the study of another focal variable ($n = 23$) or (b) addressed control variables prior to the method section but not formally in their study hypotheses ($n = 20$).
[b]Percentage of time that authors provided full descriptive information (e.g., means, standard deviations, correlations, information about gender [1 = male, 2 = female]) for a control variable.
[c]When a control is conceptual, the percentage of time that authors provided reliability information. Not included in these numbers are those cases in which it was impossible to tell if the control variable was conceptual or demographic ($n = 7$; e.g., variables labeled "planning skill," "perceived mobility," with no additional information offered).
[d]Described analyses/results with and without a control variable included.

once a study's authors select one or more control variables and include them in the write-up of their research, it is generally recommended that authors (a) include control variables in the formal hypothesis statement, (b) report standard descriptive statistics and correlations for control variables, (c) subject the control variables to the same psychometric standards (e.g., reliability) as focal variables, and (d) compare study results with and without control variables included in substantive hypothesis testing. Table 4 summarizes existing leadership research in regard to these recommendations and reveals that the leadership literature has yet to fully adopt best practices, with <1% of all studies incorporating control variables into the corresponding hypothesis and only 5% of studies reporting results with and without control variables included in their analyses. Table 4 indicates that a majority of studies do in fact provide descriptive statistics (i.e., 70% of all controls/studies overall), but this practice is far from universal. We also found that, on average, 92% of perceptual control variables are accompanied by the reporting of reliability information, but this practice is also not universal. Specifically, 10% and 13% of all transformational and transactional leadership studies, respectively, fail to report reliability estimates for conceptual controls.

What Is the Relationship Between Control Variables and Focal Leadership Variables?

Research Question 4 asked to what extent statistical control variables correlate with focal leadership variables. We explored this issue using two complementary approaches. First, we inspected a study's variable intercorrelation matrix and coded whether the control variables were significantly correlated with focal leadership constructs. As shown in Table 5,

Table 5
Qualitative Relationship Analysis Between Control Variables and Focal Leadership Domains

	Significance, % (*n*)						
Control	LMX	TFL	Transactional	Laissez-Faire	Authentic	Ethical	Shared
Gender	15% (15/103)	10% (12/117)	12% (2/17)	0% (0/4)	33% (1/3)	9% (1/11)	33% (1/3)
Follower	14% (10/72)	9% (6/65)	8% (1/12)	0% (0/3)	33% (1/3)	14% (1/7)	100% (1/1)
Leader	18% (2/11)	17% (6/35)	20% (1/5)	0% (0/1)	—	0% (0/4)	—
Gender differences	15% (3/20)	0% (0/11)	—	—	—	—	0% (0/2)
Other gender-related controls	—	0% (0/6)	—	—	—	—	—
Tenure	21% (39/185)	17% (24/138)	0% (0/15)	33% (1/3)	40% (2/5)	10% (2/20)	—
Follower job	30% (7/23)	17% (1/6)	0% (0/1)	—	—	25% (1/4)	—
Follower organizational	18% (11/60)	13% (4/31)	0% (0/3)	100% (1/1)	67% (2/3)	0% (0/5)	—
Team	60% (6/10)	11% (2/18)	—	—	0% (0/1)	0% (0/2)	—
Tenure under leader	16% (12/76)	27% (10/37)	0% (0/9)	0% (0/1)	0% (0/1)	33% (1/3)	—
Leader's job tenure	100% (1/1)	5% (1/21)	0% (0/2)	0% (0/1)	—	0% (0/2)	—
Leader's organizational tenure	14% (1/7)	25% (2/8)	—	—	—	0% (0/2)	—
Other tenure-related controls	13% (1/8)	24% (4/17)	—	—	—	0% (0/2)	—
Age	23% (23/102)	10% (10/98)	18% (3/17)	50% (1/2)	13% (1/8)	25% (2/8)	33% (3/9)
Follower	19% (15/77)	8% (5/60)	17% (2/12)	50% (1/2)	17% (1/6)	33% (2/6)	NR
Leader	56% (5/9)	17% (3/18)	50% (1/2)	—	—	0% (0/2)	—
Age differences	19% (3/16)	0% (0/6)	—	—	—	—	0% (0/1)
Organizational	—	0% (0/7)	0% (0/3)	—	0% (0/1)	—	20% (1/5)
Other age-related controls	—	29% (2/7)	—	—	0% (0/1)	—	67% (2/3)
Education	17% (8/47)	7% (3/45)	0% (0/4)	0% (0/1)	25% (1/4)	0% (0/8)	NR
Follower	24% (8/34)	6% (2/32)	0% (0/4)	0% (0/1)	25% (1/4)	0% (0/4)	NR
Leader	0% (0/8)	13% (1/8)	—	—	—	0% (0/4)	—
Education differences	0% (0/7)	0% (0/5)	—	—	—	—	—
Race	0% (0/11)	31% (11/36)	0% (0/1)	0% (0/2)	100% (2/2)	—	0% (0/1)
Follower	0% (0/4)	30% (10/30)	0% (0/1)	0% (0/2)	100% (2/2)	—	—
Leader	—	25% (1/4)	—	—	—	—	—
Racial differences	0% (0/7)	0% (0/2)	—	—	—	—	0% (0/1)
Organization Size	0% (0/5)	13% (4/34)	20% (1/5)	0% (0/1)	0% (0/1)	0% (0/2)	80% (4/5)
Team/group Size	23% (7/31)	6% (2/36)	0% (0/5)		0% (0/2)	0% (0/3)	30% (3/10)
Personality	59% (13/22)	34% (13/38)	0% (0/1)	0% (0/2)	50% (1/2)	—	50% (1/2)
Follower Big Five traits	100% (1/1)	30% (3/10)	—	—	—	—	—
Follower negative affect	50% (3/6)	0% (0/7)	0% (0/1)	—	100% (1/1)	—	100% (1/1)
Follower positive affect	83% (5/6)	75% (6/8)	—	—	—	—	—
Other personality-related controls	44% (4/9)	31% (4/13)	—	0% (0/2)	0% (0/1)	—	0% (0/1)
Follower work experience	11% (1/9)	29% (5/17)	NR	—	—	—	—
Follower work status	20% (1/5)	50% (9/18)	NR	—	—	33% (1/3)	—
Follower workload/hours	14% (1/7)	0% (0/1)	NR	—	—	100% (1/1)	—
Follower organizational level	27% (3/11)	44% (8/18)	40% (2/5)	0% (0/1)	—	—	—
Follower family-related controls	0% (0/6)	NR	NR	NR	—	—	—
Follower job function	25% (2/8)	4% (1/27)	—	0% (0/3)	—	—	—

(continued)

Table 5 (continued)

Control	Significance, % (*n*)						
	LMX	TFL	Transactional	Laissez-Faire	Authentic	Ethical	Shared
Leader span of control	0% (0/2)	43% (3/7)	67% (2/3)	—	—	0% (0/2)	—
Social desirability	0% (0/1)	33% (2/6)	—	—	—	100% (2/2)	—
Performance-related controls	50% (1/2)	20% (1/5)	—	—	—	—	—
Transformational leadership	—	—	82% (9/11)	100% (8/8)	100% (1/1)	100% (2/2)	—
Transactional leadership	—	92% (12/13)	—	75% (6/8)	—	—	—

Note: The percentage listed in this table represents the percentage of times that an effect size was designated as being significantly related to the focal variable within a primary study. This percentage was calculated per the ratio of significant effect sizes to the number of effect sizes reported within primary studies. The denominator in these percentages may not always match the number of studies listed in Table 2, because some studies included multiple ratings of the focal variable (e.g., self- and other-rated transformational leadership; transformational leadership ratings of direct manager and CEO) and/or listed only facet-level effect sizes (e.g., contingent-reward, passive management-by-exception). This ratio does not include those studies stating that the control variable had limited influence (but did not report actual correlations). When authors specifically noted that a control was insignificantly related to the focal variable, that effect size was counted as part of the ratio listed in this table. Not shown in this table are those studies that did not report effect sizes between control variables and focal variables, either because (a) they simply did not acknowledge the effects in either the correlation matrix or the manuscript text or (b) they explicitly noted in the manuscript text that controls had limited influence on study results or study results were the same with or without the inclusion of control variables. NR = not reported.

"significance" represents the percentage of times that a control variable significantly correlated with the leadership variable listed in that column.[5] To illustrate, under LMX, follower gender was correlated with LMX in 14% of reported effect sizes. Other frequently used controls were related to LMX in only 18% (follower organizational tenure), 16% (tenure under leader), 19% (follower age), and 24% (follower education) of reported effect sizes. We find a similar pattern for transformational leadership studies, with only 9% (follower gender), 17% (leader's gender), 13% (follower organizational tenure), 5% (leader's job tenure), 8% (follower age), 17% (leader's age), 13% (organizational size), and 6% (team/group size) of reported effect sizes reaching statistical significance.[6] While this qualitative information is less precise than meta-analytic summaries (reported next), it nevertheless represents an important (and previously undocumented) piece of information. This is particularly true, as the majority of control variables (77%) were not significantly correlated with focal leadership constructs.

We also explored Research Question 4 using meta-analysis. Results reported in Table 6 indicate that significant relationships exist between 5 of 16 control variables and LMX. Popular controls, such as follower organizational tenure ($\rho = .00$), follower gender ($\rho = .02$), gender differences ($\rho = -.01$), follower education ($\rho = .01$), and racial differences ($\rho = -.01$), all have negligible relationships with LMX, and all confidence intervals include zero. Moreover, the few statistical controls that were related to LMX produced effect sizes that were either relatively weak (e.g., follower age, $\rho = .04$; Cohen, 1988) or based on a small number of samples (e.g., positive affect, $k = 4$). Results for transformational and transactional leadership mirror those of LMX, with only 2 of 17 control variables relating to transformational leadership and with those 2 exhibiting weak effect sizes ($\rho = -.06$; Cohen, 1988). Finally, only three effect sizes could be computed between specific control variables and

transactional leadership, two of which produced confidence intervals that did not contain zero but were relatively small (average $\rho = -.04$).

Discussion

The ever-increasing study of leadership effects, combined with an increased emphasis on the appropriate use of statistical controls, makes the present study both timely and relevant. As little effort has been spent on learning about the use of control variables in leadership research, we drew on recommendations put forth by methodologists (see Becker et al., 2016) to systematically review 12 years of empirical research devoted to the most popular leadership topics. Overall results of our review indicate that a majority of leadership studies do not adhere to best practice recommendations for using statistical control variables. The detailed findings generating this overall assessment have a number of important implications.

Among the implications is the documentation of what factors are statistically controlled for when studying seven of the most popular leadership constructs. This is important because prior methodological work documents the average number of control variables used in a study (Atnic et al., 2012), whether such variables are perceptual or objective (Atnic et al., 2012), and the complications associated with the inclusion of controls (Becker, 2005; Breaugh, 2006; Spector & Brannick, 2011), yet no study documents what is actually being controlled for across existing leadership areas. Connected to this point is the finding that the most common statistical controls include demographic factors, such as gender, age, tenure, and education, despite repeated advice to avoid such proxy variables (e.g., Becker, 2005; Spector & Brannick, 2011). Rather than citing a single existing primary study that included/excluded one or a select few control variables, we present an overview of the leadership field that should, as a result, give researchers and reviewers more confidence in a study's methodological choices (e.g., the decision to not include leader's gender, age, etc. as control variables).

A related implication and contribution of this study is the analysis of explanations given for the statistical control of variables when studying leadership phenomena. Whereas existing research documents whether authors offer a justification (and a handful of other select categories), our study goes a step further. In comparison to previous pedagogically oriented works (Atnic et al., 2012; Becker, 2005; Carlson & Wu, 2012), our content analysis considered justification use within and between leadership areas, added more nuanced control variable justification categories, and included more journals over a greater period of time. With such an approach, we are able to offer additional insights into control variable usage within current studies of leadership. Unfortunately, these insights suggest that a relatively large percentage of leadership studies are not following best practice suggestions. Specifically, more than one-quarter of the reviewed studies failed to offer any type of explanation for their choice of control variables. When authors do offer some type of explanation, they frequently default to an assumption that the control may relate to a study focal variable, without explaining how or why this relationship might exist. In fact, <5% of all leadership studies including statistical controls offer a theoretical explanation for control variable usage. Extrapolating from these findings, we observe that the overabundance of demographic control variables may be explained by the poor "why" decisions made when statistically controlling for a variable.

Another implication of this study is that we assessed whether current research is following "essential recommendations" put forth to enhance the quality of empirical research (Becker

Table 6
Meta-Analysis of the Relations Between Leadership Constructs and Potential Control Variables

	Leader-Member Exchange						Transformational Leadership						Transactional Leadership					
Control Variable	*k*	*n*	ρ	SD_ρ	95% CI	80% CV	*k*	*n*	ρ	SD_ρ	95% CI	80% CV	*k*	*n*	ρ	SD_ρ	95% CI	80% CV
Follower gender[a]	56	15,372	.02	.06	[−.01, .04]	[−.05, .09]	41	14,646	−.01	.06	[−.04, .02]	[−.08, .06]	8	1,876	−.06	.00[b]	[−.11, −.01]	—
Leader gender[a]	9	2,746	.02	.08	[−.05, .09]	[−.08, .11]	20	3,261	−.04	.07	[−.10, .03]	[−.13, .05]	—	—	—	—	—	—
Gender differences	14	2,792	−.01	.07	[−.08, .05]	[−.10, .07]	6	559	.05	.05	[−.05, .16]	[−.01, .12]	—	—	—	—	—	—
Follower job tenure	15	3,884	−.03	.09	[−.09, .03]	[−.14, .09]	6	1,584	−.07	.04	[−.14, −.00]	[−.12, −.01]	—	—	—	—	—	—
Follower organizational tenure	43	12,646	.00	.04	[−.02, .03]	[−.05, .06]	27	10,481	−.04	.07	[−.07, .−01]	[−.12, .05]	—	—	—	—	—	—
Tenure under leader	42	9,400	.06	.06	[.03, .09]	[−.02, .14]	21	7,811	−.01	.07	[−.05, .04]	[−.10, .08]	4	2,972	−.02	.00[b]	[−.04, −.01]	—
Leader job tenure	—	—	—	—	—	—	13	1,618	−.04	.03	[−.10, .02]	[−.08, .00]	—	—	—	—	—	—
Leader organizational tenure	5	1,510	.02	.13	[−.13, .17]	[−.14, .19]	7	1,041	−.05	.09	[−.20, .10]	[−.17, .07]	—	—	—	—	—	—
Follower age	58	15,031	.04	.08	[.01, .07]	[−.06, 14]	34	10,429	.01	.08	[−.03, .05]	[−.09, .11]	6	1,633	.01	.06	[−.12, .14]	[−.06, .09]
Leader age	8	2,144	.06	.13	[−.07, .18]	[−.11, .23]	15	2,801	.00	.10	[−.08, .09]	[−.12, .13]	—	—	—	—	—	—
Age differences	12	2,526	−.03	.08	[−.09, .03]	[−.13, .07]	4	419	−.01	.00[b]	[−.11, .10]	—	—	—	—	—	—	—
Organizational age	—	—	—	—	—	—	7	571	.05	.00[b]	[−.03, .12]	—	—	—	—	—	—	—
Follower education	27	8,736	.01	.07	[−.03, .05]	[−.08, .10]	19	5,853	.05	.06	[−.01, .10]	[−.03, .12]	—	—	—	—	—	—
Leader education	6	1,742	−.05	.00[b]	[−.08, −.01]	—	6	934	.03	.12	[−.13, .19]	[−.13, .19]	—	—	—	—	—	—
Racial differences	6	1,002	−.01	.00[b]	[−.08, .05]	—	—	—	—	—	—	—	—	—	—	—	—	—
Follower negative affect	4	1,080	−.16	.09	[−.28, −.04]	[−.27, −.05]	3	860	.02	.00[b]	[−.00, .04]	—	—	—	—	—	—	—
Follower positive affect	4	1,080	.23	.00[b]	[.21, .25]	—	4	1,185	.18	−.01	[−.01, .36]	[−.02, .37]	—	—	—	—	—	—
Follower work experience	8	1,748	.04	.00[b]	[−.00, .09]	—	5	1,110	.03	−.13	[−.13, .18]	[−.13, .17]	—	—	—	—	—	—

Note: Effects sizes and 95% confidence intervals (95% CIs) represent corrected reliability estimates. CV = credibility interval.

[a]Scored such that males are represented by larger numbers.

[b]For these analyses, the estimated average sampling error exceeds the variance of ρ, resulting in a negative residual variance. As a standard deviation cannot be computed from a negative variance, Hunter and Schmidt (2004) assert that such instances should be regarded as having zero variance.

et al., 2016: 157). Whereas we drew heavily from Becker and colleagues' (2016) recent guidance for control variable usage, their recommendations are a conglomerate of previous work, some of which dates back to the 1970s (e.g., Meehl, 1970). Even as methodologists repeatedly raise these issues, it is clear many authors, editors, and reviewers have not heeded their recommendations (Green, Tonidandel, & Cortina, 2016). In fact, we found that nearly 30% of control variables were missing important descriptive information: anything from means and standard deviations to reliability information and explanation of what a variable actually represents (e.g., something categorical or something conceptual). Also rare is the recommended reporting habit to explicitly include control variables in one's hypotheses even though subsequent tests do in fact focus on conditional relationships (i.e., control variables are part of the statistical analyses). One final concern in relation to the quality of empirical research surrounds the reporting of results with and without control variables. In spite of repeated calls among methodologists (Aguinis & Vandenberg, 2014; Becker, 2005), we found that only 5% of studies report comparisons of study results with and without control variables in the analyses. This is particularly surprising (to us at least) given that Becker initially recommended this practice a decade ago and that Becker's article has been cited >600 times according to Google Scholar.

One final implication of this study is the analysis of existing relationships between control variables and leadership constructs (i.e., significance, direction). Notably, a qualitative examination indicated that fewer than one-fifth of all control variables were related to the leadership variables, and a meta-analytic examination of the relationships between the most commonly used controls and LMX, transformational, and transactional leadership found that the average overall effect size was |.04|. Given previous calls for researchers to more explicitly describe expected relationships between control variables and focal variables (e.g., Spector & Brannick, 2011), these findings provide a key piece of information to the leadership literature. Specifically, these findings may help prevent the inclusion of unnecessary controls that not only burn degrees of freedom but also open up results to other statistical complications (e.g., empirical suppression) while muddying the interpretation of results. Insofar as our systematic review documents the weak effect sizes found between leadership constructs and commonly used statistical controls, we hope that this finding draws attention to the need for scholars to appreciate the critical complications associated with statistical control and to take a conservative approach to control variable usage. To this point, good study design only increases the likelihood of generalizable and valid findings (Aguinis & Vandenberg, 2014).

Moving Forward: Best Practice Recommendations for Control Variables

In the majority of studies included in this review, researchers included demographic variables in an attempt to purify their results of potential confounds and/or alternative third variable explanations. This is an unfortunate finding given that demographic variables are rarely the variables of interest but rather serve as convenient proxies used in place of conceptually meaningful control variables. Moreover, given the interpretation difficulties associated with control variables usage (e.g., Becker et al., 2016; Breaugh, 2006; Spector & Brannick, 2011), one could easily question whether our understanding of leadership phenomena is being unduly influenced by the inappropriate use of statistical controls. Considering this possibility, we call for a moratorium on the use of all proxy variables while simultaneously calling for the discontinuation of using demographic control variables unless there is a clear and

compelling theoretical rationale for doing so. Instead, we challenge leadership researchers to identify specific, conceptually meaningful control variables (and corresponding measures) that match the study's purpose, while also urging journal editors and reviewers to question and even return submitted manuscripts that do not follow best practice recommendations.

Our repeated call to focus on theoretically relevant controls while avoiding demographic proxy variables also allows us to make additional observations in regard to potentially appropriate control variables for leadership research. Existing advice makes it clear that proxy (e.g., demographic) variables should not be treated as statistical controls, but the question of which variables could be included is less clear. We believe that the answer to this question depends largely on the existence of theory and, in particular, a clear understanding of a leadership construct's nomological network. Consider transformational leadership theory and its fundamental theoretical assumption that it augments the beneficial effects of transactional leadership (Bass, 1985). As a result of this theoretical connectedness, studies of transformational leadership should control for transactional leadership behaviors in an effort to account for alternative theoretical explanations. To be sure, the "augmentation effect" is often discussed in the literature (e.g., Avolio, 1999; Bass, Avolio, Jung, & Berson, 2003; Cole, Bedeian, & Bruch, 2011), but our findings show that it is rarely tested. Likewise, a review of the leadership literature (Dinh et al., 2014) suggests a relative explosion of new leadership theories and measures. As this new wave of "emerging" leadership perspectives suggests that there is more to leadership than forming high-quality relationships or inspiring followers, it seems theoretically appropriate and methodologically necessary to account for already-existing and heavily researched leadership constructs (i.e., TFL and LMX) when testing the uniqueness of these new leadership constructs.

If a clear theoretical basis is not available for utilizing a control variable yet the author feels that it is nevertheless important to incorporate it into one's analyses, its adoption requires a rich explanation of how it might bias predicted relationships (Becker et al., 2016). One such variable in the leadership area is target-specific affect. Notably, D. J. Brown and Keeping (2005) demonstrated that a rater's (i.e., follower's) liking of a target individual (i.e., leader) influenced the measurement of transformational leadership as well as its substantive relationships with outcomes. As such, controlling for target-specific affect or liking appears justified in the study of transformational leadership because it can contaminate followers' responses to survey instruments. In this regard, it also seems reasonable to assume that, to the extent that target-specific liking is a salient source of bias, it may also contaminate follower-generated responses to any number of leadership instruments.

Before continuing, we think that it is also important to acknowledge that not all methodologists take the stance that less is more. Antonakis, Bendahan, Jacquart, and Lalive (2014), for example, discuss how the inclusion of control variables can address endogeneity concerns. These authors note that including *more* theoretically relevant controls can help authors establish causal claims in nonexperimental research designs. We do not necessarily endorse Antonakis and colleagues' conclusion that more control variables are better, but we wholeheartedly support their assertion that having the right controls is essential. Thus, whereas many methodologists (e.g., Becker et al., 2016; Carlson & Wu, 2012) suggest that less is more when it comes to statistical control, perhaps a better recommendation—and the one that we put forth—is that less of the nontheoretical type is more.

Building on the basic notion of excluding and including control variables based on a sound theoretical foundation leads us to make additional recommendations about commonly used control variable *justifications* identified in our review. We summarize these recommendations

in Table 7. First, we submit that including a particular control variable simply because others have done so is not a sufficient reason for incorporating it into one's analytical model. In other words, a study that states something to the effect of "Consistent with prior studies of LMX, we controlled for dyadic tenure" intimates that the authors have not thoughtfully selected their controls.[7] For journal gatekeepers, this justification type should raise immediate red flags. A reference to the debunked fallacy of "conservative, rigorous, or stringent" tests (Carlson & Wu, 2012; Meehl, 1971; Spector & Brannick, 2011) represents another commonly used justification that lacks a sound basis and should likewise raise red flags that send manuscripts back to submitting authors. Third, we contend that simply stating that a control variable has a potential relationship with a leadership variable or that previous research found an empirical relationship between a control variable and a leadership variable is not enough to warrant its inclusion; instead, researchers need to include a full account of prior evidence (e.g., correlational or experimental) in addition to explaining how and why these control variables relate to what they are studying and how including/excluding such variables will impact their research.

Three justifications that many would agree support control variable use include accounting for method bias (and other forms of contamination), alternative explanations, and assessing incremental validity (Carlson & Wu, 2013). Whereas all three of these explanations have the potential to satisfy methodologists' concerns if implemented correctly and presented in a detailed theoretical manner, proposing method bias or alternative explanations as a justification for control variable use presents additional challenges for proper implementation and interpretation of findings. When method bias and other forms of contamination are a concern, the challenge is that researchers need to understand exactly why the bias is occurring and whether it is affecting the predictor or criterion (or both). With that understanding, researchers then need to identify a meaningful control variable that captures only the bias or contamination (i.e., it is independent of the covariance between the predictor and criterion). This is admittedly a tough task and one that our findings suggest is not widely understood.

As pointed out by an anonymous reviewer, it is also important to recognize that in hierarchical multiple regression (a frequently used approach when testing hypotheses), control variables entered into the first block of analysis partial shared variance out of the predictor and not the criterion. In other words, if the control variable is selected because researchers believe that it relates to the criterion, the "control effect" that actually occurs is between the control variable and the predictor; the impact of such "control effects" varies as a result of the extent to which the two variables (i.e., control and predictor, not control and criterion) are correlated. Hence, a potential challenge when controlling for alternative explanations is related to the partialing of shared variance between the control variable(s) and the predictor. Consider a scenario in which the control variable and predictor are capable of explaining a roughly equal amount of variance in the criterion. Because of the shared variance, the mathematical partialing of the variance depends on the extent of covariation between the control(s) and predictor as well as the rank order of the bivariate (zero-order) relationships between these variables and the criterion. In such a scenario, the regression coefficients can differ substantively from their respective bivariate correlations, making interpretations of regression coefficients quite challenging. It is therefore important for scholars to recognize that their job is not complete even after using a strong theoretical explanation to justify their control variable decisions.

One additional justification that can satisfy methodological concerns is best described as a process of analysis. Inspection of study data as a stand-alone justification and/or analytic approach (i.e., controlling after results are known) is not defensible, but if included in a sequence of methodological choices that begin with a deep theoretical understanding of the leadership

Table 7
Appropriate, Inappropriate, and Best Practice Recommendations for Justifying Statistical Controls

Typical Justification	Assessment	Recommendation
Imitating other works/ authors	A stand-alone statement such as "Following other researchers, we controlled for . . ." is not enough to warrant inclusion.	Citing the use of certain control variables in previous research is not enough to warrant inclusion unless authors also explain how and why such variables relate to their research questions. In such cases, citing the empirical findings of previous research that also used such variables can help primary researchers make ultimate inclusion or exclusion decisions.
Robustness	Suggesting that including controls offers a more "rigorous," "conservative," or "stringent" test is not enough to warrant inclusion.	Statements such as these are fallacious and should not be used going forward.
Potential relationships	Stating that control variables could potentially relate to focal study variables is not enough to warrant inclusion without offering more explanation.	Suggesting that a potential relationship between a control variable and a study focal variable is justifiable only if the authors describe how and why such relationships likely exist. Best practice includes such descriptions in the hypothesis-building section.
Previous empirical evidence	Offering evidence that control variables empirically relate to focal study variables in previous research is not enough to warrant inclusion.	Citing previous empirical evidence that control variables relate to study focal variables is important, but authors must explain how and why such relationships exist. Citing previous findings and explaining how and why such findings exist, explaining why including such variables is important, and indicating how such variables might influence their research is best done in the hypothesis-building section.
Common method bias	Suggesting that controls "partial out method variance" is not enough to warrant inclusion.	Purely methodological explanations need to be explicit in study deficiencies, explaining why such characteristics (e.g., bias) exist, how and why deficiencies could affect study relationships, and how the use of statistical controls addresses such deficiencies.
Possible alternative explanations	Blanket statements such as "We control for X, Y, and Z to eliminate alternative explanations" is not enough to warrant inclusion.	Authors need to explain in what way controls serve as an alternative explanation, including how controls relate to a study's focal variables, how including them eliminates these alternative explanations, and why eliminating alternative explanations is important.
Incremental validity	Attempting to demonstrate incremental validity is a sound practice if implemented correctly.	Authors should explain how and why controls relate to focal criteria and why it is important to demonstrate focal predictors explain unique variance above controls.
Analyzing study (i.e., one's own) data	Simply referencing empirical relationships found in study data is not enough to warrant inclusion.	Researchers should analyze study data to find insignificant controls, but there must be a theoretical reason for considering a variable as a control prior to empirical examination. If theory suggests a possible relationship between a control and a focal variable, researchers should examine the empirical relationship. If no relationship exists, excluding potential controls is justifiable but should be described in the manuscript. Analyzing study data to find controls that enhance or hurt hypothesized relationships is never acceptable.
Multiple/mixing justifications	Taking a combination of steps is best practice.	Best practice recommendations for control variable use begins with a strong theoretical explanation of how and why potential control variables relate to study focal variables. This should include both prior theoretical and empirical work. Researchers should explain what purpose including such variables (e.g., demonstrating incremental validity) serves and include this description in the hypothesis development section. Researchers should also analyze study data and exclude impotent variables. Each of these actions should be explicitly described in the manuscript, including reporting descriptive statistics for included and excluded controls.

construct's nomological network, such analysis helps strengthen one's ultimate inclusion or exclusion decisions. In fact, we propose that best practice justifications describe how and why control variables relate to focal study variables (i.e., offer a theoretical explanation), indicate why it is important to include such controls (e.g., incremental validity), quantitatively analyze relationships (e.g., correlations) between controls and focal study variables, and compare study results with and without controls before making ultimate inclusion or exclusion decisions.

Study Limitations

Whereas we included studies from more than a decade of leadership research in top management and applied psychology journals, our review is not without potential limitations. One possible limitation is that some coded studies used the same justification for the inclusion of two or more control variables and/or the same justification for the inclusion of control variables across multiple leadership areas (e.g., in the study of transformational and transactional leadership). Consequently, these explanations were weighted more heavily when calculating overall percentages within studies and across leadership domains than other explanations that were used to justify the inclusion of a singular control or in the study of a singular leadership construct. We also note that some control variable justifications did not explicitly reference a relationship with a leadership construct; instead, the study offered a generic statement (e.g., controls could relate to study variables) without reference to a particular focal leadership variable. We included these instances in our analyses because it was impossible to determine if a "nonleadership" variable was the only reason why the control was included and because ultimately control variables were statistically controlled for when analyzing leadership styles. A quick analysis of study results without these studies suggests an equivalent pattern of results. A third possible concern that we wish to acknowledge is that the seven leadership constructs may assume the role of both exogenous and endogenous variables. We attempted to discern differences between the justifications and variables used across these roles, but because leadership variables predominately serve as predictor variables and infrequently serve as criterion, there were not enough combinations to allow for meaningful comparisons.

Conclusion

The use of statistical control as a methodological tool is prevalent in field research on leadership. As such, it is important to understand what variables are being controlled for, why such variables are controlled, and how those variables relate to popular leadership topics. Our study provides evidence that a majority of control variables relate weakly—theoretically and empirically—to commonly researched leadership topics. Our hope is this study serves as a valuable resource, making control variable assumptions and potential problems more accessible and salient to leadership researchers and perhaps social science scholars in general, including those studying macrolevel management topics who utilize, on average, 10 control variables per study (Atinc et al., 2012: 67). Toward this end, we encourage future research to reconsider using many of the variables found in this review in favor of more theoretically relevant factors.

Notes

1. Scholars use different terms when describing these third variable effects, including "control" and "nuisance" variables. For some researchers, statistical control variables reference correlates of substantive interest—those theoretically related to focal variables (Becker et al., 2016). These scholars distinguish such variables from other variables, referred to as nuisance variables, that contaminate or artificially inflate relationships of interest without a theoretical reason for doing so. This slight distinction on conceptual grounds likely creates different expectations about the intended role of third variables, but they are identical statistically (MacKinnon et al., 2000). For the purposes of our research, we use the term *control variable* as a catch-all term because any third variable becomes a "statistical control" when entered into a mathematical model.

2. Each of these facets of leadership is generally accepted as a component of leadership, yet there is no universal agreement on how such facets should be grouped. We include charisma under transformational leadership even though some researchers suggest that charisma and transformational leadership are "distinct but overlapping processes" (Yukl, 1999: 299). While acknowledging such views, we ultimately grouped charisma with other facets of transformational leadership because this approach is consistent with several seminal meta-analyses on transformational leadership (e.g., Eagly, Johannesen-Schmidt, & van Engen, 2003; Judge & Piccolo, 2004; Lowe, Kroeck, & Sivasubramaniam, 1996) and has also been recognized as the "flagship" approach by notable leadership scholars (see Antonakis, 2012: 257). We also note that passive management-by-exception is frequently included under transactional leadership. We, however, grouped passive management-by-exception with laissez-faire leadership for several reasons, including "resemblance" between the two, including conceptual and empirical overlap (Judge & Piccolo, 2004: 756), and because passive management-by-exception shows inverse relationships with other transactional leadership facets—meaning that any attempt to meta-analyze transactional leadership effect sizes with popular controls would potentially zero out any existing relationships if grouped together. Our approach, while not universally accepted, is consistent with other scholars (e.g., Bono et al., 2012) and even notable meta-analyses (e.g., Bono & Judge, 2004; DeRue et al., 2011).

3. For a number of control variables and several leadership styles (i.e., laissez-faire, authentic, ethical, shared), meta-analytic analysis was not feasible, as primary researchers used incompatible response options (e.g., different racial categories) or there were simply not enough samples to result in any type of meaningful analysis (e.g., leaders' gender, $k = 2$ for transactional leadership). We also note that because leadership variables operate uniquely at the individual and group levels, meta-analytic effect sizes reported in Table 6 include only individual-level effect sizes.

4. As a supplement to the analysis reported in Table 3, we also analyzed control variables across specific control variables (e.g., leader gender, organizational size) and across the 12-year period to determine if authors offered more theoretical justifications across specific variables and/or to assess if practices have changed over time. Unfortunately, these results indicate poor control variable practices span across various control variables and across the last 12 years. In particular, no individual control variable included a theoretical justification in >5% of studies, and no year, including 2013 and 2014, included theoretical justifications in >5% of studies. While explanations do not appear to be improving overall, one potentially positive trend found in control variable usage is the increase in control variable "analysis" over the past 6 years. That is, more authors are assessing the relationship that control variables hold with focal variables prior to ultimate inclusion or exclusion (average analysis = 19% of studies for 2009-2014 vs. 3% of studies for 2003-2008).

5. As seen in Table 5, we qualitatively document the significance between focal leadership styles and included control variables. Whereas we feel that this is an important piece of information to get a general sense of the types of relationships occurring in existing research, we note that significance level is oftentimes a consequence of sample size.

6. Closer examination of Table 5 indicates that there are some control variables under each leadership domain that correlate more frequently with the focal leadership variable, but these results were typically based on <5 studies; hence, the generalizability of these results is undeterminable at the present time.

7. The quotation in this sentence comes directly from one of the studies included in our analysis. Because this type of justification is frequently offered with leadership and management/applied psychology research in general, we did not feel it was necessary to criticize individual authors.

References

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.

Antonakis, J. 2012. Transformational and charismatic leadership. In D. V. Day & J. Antonakis (Eds.), *The nature of leadership*: 256-288. Thousand Oaks, CA: Sage.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. 2014. Causality and endogeneity: Problems and solutions. In D.V. Day (Ed.), *The Oxford handbook of leadership and organizations*: 93-117. New York: Oxford University Press.

Atinc, G., Simmering, M. J., & Kroll, M. J. 2012. Control variable use and reporting in macro and micro management research. *Organization Research Methods*, 15: 57-74.

Avolio, B. J. 1999. *Full leadership development: Building the vital forces in organizations*. Thousand Oaks, CA: Sage.

Avolio, B. J., & Bass, B. M. 1990. *The full range leadership development programs: Basic and advanced manuals*. Binghamton, NY: Bass, Avolio, & Associates.

Bacha, E. 2014. The relationship between transformational leadership, task performance, and job characteristics. *Journal of Management Development*, 33: 410-420.

Bacharach, S. B. 1989. Organizational theories: Some criteria for evaluation. *Academy of Management Review*, 14: 496-515.

Bass, B. M. 1985. *Leadership and performance beyond expectations*. New York: Free Press.

Bass, B. M., Avolio, B. J., Jung, D. I., & Berson, Y. 2003. Predicting unit performance by assessing transformational and transactional leadership. *Journal of Applied Psychology*, 88: 207-218.

Becker, T. E. 2005. Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods*, 8: 274-289.

Becker, T. E., Atinc, G., Breaugh, J. A., Carlson, K. D., Edwards, J. R., & Spector, P. E. 2016. Statistical control in correlational studies: 10 essential research recommendations for organizational researchers. *Journal of Organizational Behavior*, 37: 157-167.

Bernerth, J. B., & Aguinis, H. 2016. A critical review and best-practice recommendations for control variable usage. *Personnel Psychology*, 69: 229-283.

Bernerth, J. B., Armenakis, A. A., Feild, H. S., Giles, W. F., & Walker, H. J. 2007. Leader-member social exchange (LMSX): Development and validation of a scale. *Journal of Organizational Behavior*, 28: 979-1003.

Bono, J. E., Hooper, A., & Yoon, D. J. 2012. The impact of personality on leadership ratings. *The Leadership Quarterly*, 23: 132-145.

Bono, J. E., & Judge, T. A. 2004. Personality and transformational and transactional leadership: A meta-analysis. *Journal of Applied Psychology*, 89: 901-910.

Breaugh, J. A. 2006. Rethinking the control of nuisance variables in theory testing. *Journal of Business and Psychology*, 20: 429-443.

Breaugh, J. A. 2008. Important considerations in using statistical procedures to control for nuisance variables in non-experimental studies. *Human Resource Management Review*, 18: 282-293.

Brown, D. J., & Keeping, L. M. 2005. Elaborating the construct of transformational leadership: The role of affect. *The Leadership Quarterly*, 16: 245-272.

Brown, M. E., & Treviño, L. K. 2006. Ethical leadership: A review and future directions. *The Leadership Quarterly*, 17: 595-616.

Carlson, K. D., & Wu, J. 2012. The illusion of statistical control: Control variable practice in management research. *Organizational Research Methods*, 15: 413-435.

Carson, J. B., Tesluk, P. E., & Marrone, J. A. 2007. Shared leadership in teams: An investigation of antecedent conditions and performance. *Academy of Management Journal*, 50: 1217-1234.

Charbonnier-Voirin, A., Akremi, A. E., & Vandenberghe, C. 2010. A multilevel model of transformational leadership and adaptive performance and the moderating role of climate for innovation. *Group & Organization Management*, 35: 699-726.

Choudhary, A. I., Akhtar, S. A., & Zaheer, A. 2013. Impact of transformational and servant leadership on organizational performance: A comparative analysis. *Journal of Business Ethics*, 116: 433-440.

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cole, M. S., Bedeian, A. G., & Bruch, H. 2011. Linking leader behavior and leadership consensus to team performance: Integrating direct consensus and dispersion models of group composition. *The Leadership Quarterly*, 22: 383-398.

Dahling, J. J., Chau, S. L., & O'Malley, A. 2012. Correlates and consequences of feedback orientation in organizations. *Journal of Management*, 38: 531-546.

Day, D. V., & Antonakis, J. 2011. *The nature of leadership*. Thousand Oaks, CA: Sage.

DeRue, S. D., Nahrgang, J. D., Wellman, N., & Humphrey, S. E. 2011. Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. *Personnel Psychology*, 64: 7-52.

Dinh, J. E., Lord, R. G., Gardner, W. L., Meuser, J. D., Liden, R. C., & Hu, J. 2014. Leadership theory and research in the new millennium: Current theoretical trends and changing perspectives. *The Leadership Quarterly*, 25: 36-62.

Dulac, T., Coyle-Shapiro, J., Henderson, D. J., & Wayne, S. J. 2008. Not all responses to breach are the same: The interconnection of social exchange and psychological contract processes in organizations. *Academy of Management Journal*, 51: 1079-1098.

Duriau, V. J., Reger, R. K., & Pfaffer, M. D. 2007. A content analysis of the content analysis literature in organization studies: Research themes, data sources, and methodological refinements. *Organizational Research Methods*, 10: 5-34.

Eagly, A. H., Johannesen-Schmidt, M. C., & van Engen, M. L. 2003. Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men. *Psychological Bulletin*, 129: 569-591.

Edwards, J. R. 2008. To prosper, organizational psychology should . . . overcome methodological barriers to progress. *Journal of Organizational Behavior*, 29: 469-491.

Gardner, W. L., Lowe, K. B., Moss, T. W., Mahoney, K. T., & Cogliser, C. C. 2010. Scholarly leadership of the study of leadership: A review of *The Leadership Quarterly*'s second decade, 2000-2009. *The Leadership Quarterly*, 21: 922-958.

George, W. 2003. *Authentic leadership: Rediscovering the secrets to creating lasting value*. San Francisco: Jossey-Bass.

Graen, G. B., & Scandura, T. A. 1987. Toward a psychology of dyadic organizing. *Research in Organizational Behavior*, 9: 175-208.

Green, J. P., Tonidandel, S., & Cortina, J. M. 2016. Getting through the gate: Statistical and methodological issues raised in the reviewing process. *Organizational Research Methods*, 19: 402-432.

Hiller, N. J., DeChurch, L. A., Murase, T., & Doty, D. 2011. Searching for outcomes of leadership: A 25-year review. *Journal of Management*. 37: 1137-1177.

Hunter, J. E., & Schmidt, F. L. 2004. *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage.

Judge, T. A., & Piccolo, R. F. 2004. Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, 89: 755-768.

Liden, R. C., Erdogan, B., Wayne, S. J., & Sparrowe, R. T. 2006. Leader-member exchange, differentiations, and task interdependence: Implications for individual and group performance. *Journal of Organizational Behavior*, 27: 723-746.

Lowe, K. B., Kroeck, K. G., & Sivasubramaniam, N. 1996. Effectiveness correlates of transformational and transactional leadership: A meta-analytic review of the MLQ literature. *The Leadership Quarterly*, 7: 385-425.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. 2000. Equivalence of the mediation, confounding, and suppression effect. *Prevention Science*, 1: 173-181.

McMurray, A. J., Islam, M. M., Sarros, J. C., & Pirola-Merlo, A. 2012. The impact of leadership on workgroup climate and performance in a non-profit organization. *Leadership & Organization Development Journal*, 33: 522-549.

Meehl, P. E. 1970. Nuisance variables and the ex post facto design. In M. Radner & S. Winokur (Eds.), *Analyses of theories and methods of physics and psychology*: 373-402. Minneapolis: University of Minnesota Press.

Morrison, E. 2010. OB in AMJ: What is hot and what is not? *Academy of Management Journal*, 53: 932-936.

Pastor, J. C., Mayo, M., & Shamir, B. 2007. Adding fuel to fire: The impact of followers' arousal on ratings of charisma. *Journal of Applied Psychology*, 92: 1584-1596.

Sears, G., & Holmvall, C. 2010. The joint influence of supervisor and subordinate emotional intelligence on leader-member exchange. *Journal of Business Psychology*, 25: 593-605.

Sekiguchi, T., Burton, J. P., & Sablynski, C. J. 2008. The role of job embeddedness on employee performance: The interactive effects with leader-member exchange and organization-based self-esteem. *Personnel Psychology*, 61: 761-792.

Sluss, M. D., & Thompson, B. S. 2012. Socializing the newcomer: The mediating role of leader-member exchange. *Organizational Behavior and Human Decision Processes*, 119: 114-125.

Spector, P. E., & Brannick, M. T. 2011. Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods*, 14: 287-305.

Sutton, R. I., & Staw, B. M. 1995. What theory is not. *Administrative Science Quarterly*, 40: 371-384.

Wang, D., Waldman, D. A., & Zhang, Z. 2014. A meta-analysis of shared leadership and team effectiveness. *Journal of Applied Psychology*, 99: 181-198.

Xu, E., Huang, X., Lam, C. K., & Miao, Q. 2012. Abusive supervision and work behaviors: The mediating role of LMX. *Journal of Organizational Behavior*, 33: 531-543.

Yukl, G. 1999. An evaluation of conceptual weaknesses in transformational and charismatic leadership theories. *The Leadership Quarterly*, 10: 285-305.

Zaccaro, S. J., & Horn, Z. N. J. 2003. Leadership theory and practice: Fostering an effective symbiosis. *The Leadership Quarterly*, 14: 769-806.

Zickar, M. J., & Highhouse, S. 2001. Measuring prestige of journals in industrial-organizational psychology. *The Industrial-Organizational Psychologist*, 38: 29-36.