Journal of Management
Vol. XX No. X, Month XXXX 1–26
DOI: 10.1177/0149206316687295
© The Author(s) 2017
Reprints and permissions:
sagepub.com/journalsPermissions.nav

A New Perspective on Method Variance: A Measure-Centric Approach

Paul E. Spector
University of South Florida
Christopher C. Rosen
University of Arkansas
Hettie A. Richardson
Texas Christian University
Larry J. Williams
University of Nebraska–Lincoln
Russell E. Johnson
Michigan State University

A widespread methodological concern in the organizational literature is the possibility that observed results are due to the influence of common-method variance or mono-method bias. This concern is based on a conception of method variance as being produced by the nature of the method itself, and therefore, variables assessed with the same method would share common-method variance that inflates observed correlations. In this paper, we argue for a more complex view of method variance that consists of multiple sources that affect each measured variable in a potentially unique way. Shared sources among measures (common-method variance) act to inflate correlations, whereas unshared sources (uncommon-method variance) act to attenuate correlations. Two empirical examples, one from a simulation study and the other from a single-source survey, are presented to illustrate the complex action of multiple sources of method variance. A five-step approach is suggested whereby a theory of the measure is generated for each measured variable that would inform strategies to control for method variance by assessing and modeling the actions of identified method variance sources.

Keywords: *common-method variance; uncommon-method variance; method bias; psychometrics; construct validity*

Acknowledgments: This paper is based in part on a 2013 Southern Management Association Conference Ask the Experts session, "Advice for Identifying and Remedying the Effects of Common Method Variance." The authors would like to thank Eric Yochum for his help in preparing this manuscript.

Corresponding author: Paul E. Spector, Department of Psychology, PCD 4118, University of South Florida, Tampa, FL 33620, USA.

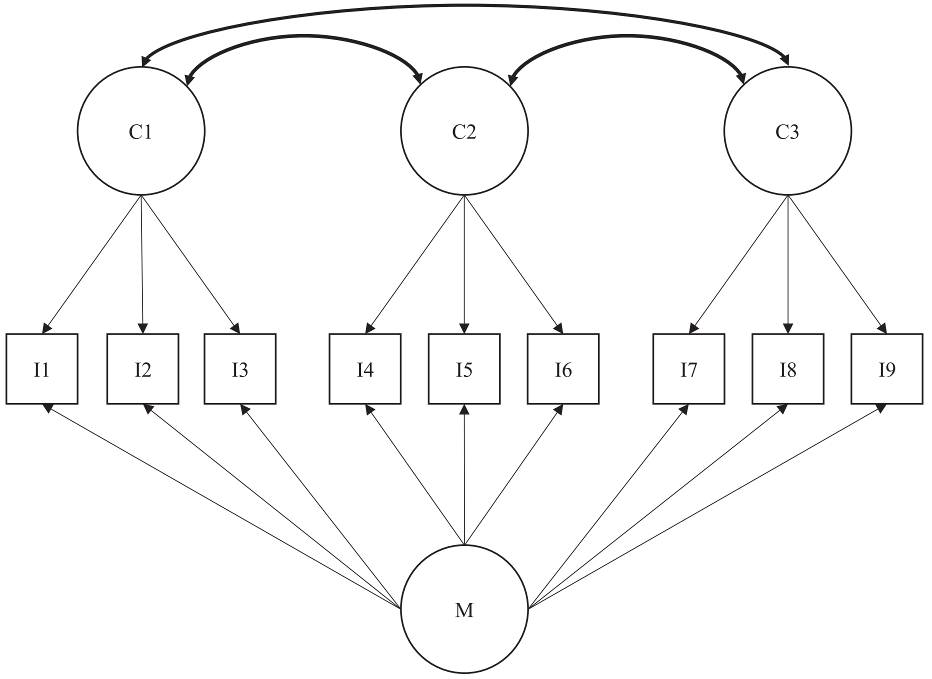
E-mail: pspector@usf.edu

There are few methodological problems that are more discussed among organizational researchers than the possibility that observed relationships among measures are distorted by the impact of method variance. The traditional definition of method variance, as noted by Campbell and Fiske (1959), is that method variance is variation in observations that is due to the method used rather than the constructs of interest; that is, method variance resides in the method. If two measures use the same method, shared- or common-method variance (CMV) components can inflate the magnitude of relationships (e.g., correlations) among variables. Less frequently considered is that unshared- or *uncommon*-method variance (UMV) can attenuate correlations (Williams & Brown, 1994). Given that this issue is generally not discussed (or explicitly addressed) in primary studies, this possibility seems rarely to be of concern.

The focus in almost all discussions of the method variance issue has been on the method used, most notably, self-reports. It has been noted that variables can vary in the amount of method variance they contain (Richardson, Simmering, & Sturman, 2009, refer to this as the congeneric case), and it has been noted that the construct as well as the method are important (Fiske, 1987b; Spector & Brannick, 1995). We go a step farther in this paper, arguing for a new perspective on method variance. Rather than approaching method variance as something that arises due to the nature of a method, we suggest that method variance represents extraneous and unintended systematic influences on a measured variable, some of which might be shared with other measured variables (viz., CMV) and some of which is not (viz., UMV). This new perspective opens up new approaches to dealing with method variance that consider both CMV and UMV. We argue that these new approaches have the potential to produce more accurate estimates of construct relationships than those that exclusively focus on method as the source of method variance. The failure to include explicit measures of method variance in an investigation results in model misspecification at the measurement level. Just as model misspecification at the construct level can lead to erroneous inference, so too can measurement misspecification.

In this paper we make a case for considering method variance to be a more complex problem requiring far more extensive strategies than have been previously advanced. To this end, we address five major purposes. First, we argue that method variance is best approached not from the perspective of the method used but at the level of the individual measured variable. As we will note, the nature of the construct is as important as the method used to assess it. Failure to consider the sources of method variance for all measures in a study will likely result in misspecification of the study's underlying measurement model that can lead to incorrect estimates of construct interrelationships. Second, we discuss the necessity of considering both CMV and UMV as factors that can distort observed relationships in opposing directions. In any given investigation it is important to consider the net effect of both CMV and UMV in order to determine whether observed relationships among theoretical constructs are inflated or deflated, or whether CMV and UMV cancel one another out. This must be done at the level of the observed relationship as each variable pair will have its own mix of CMV and UMV. Third, we provide two empirical examples to illustrate the complex action of multiple sources of method variance. Fourth, we present a strategy by which researchers can identify potential sources of method variance for specific measures. This procedure involves incorporating the specification of method variance sources into a theory of a measure that should routinely be part of measure construction and validation. Finally, we argue

Figure 1
Traditional Common-Method Variance Model

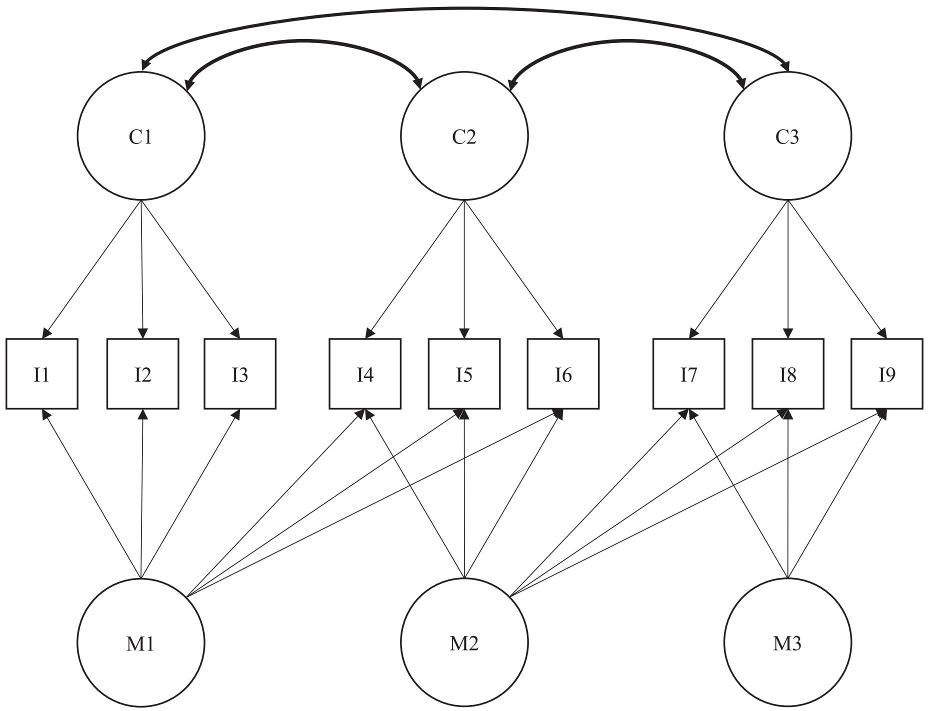


that method variance control strategies should be selected for empirical studies based on measure theories rather than the generic approach that is currently popular in the organizational sciences. Such approaches need to consider not only CMV but UMV as well to provide accurate estimates of relationship size and to avoid potential misspecification of tested models.

Method Variance as Measurement Model Misspecification

Method variance is generally defined as *systematic variation in an observed variable due to the method used*. Figure 1 uses standard structural equation modeling notation to illustrate the connection between latent constructs (circles) and observed indicators (squares). The *C*s (i.e., C1–C3) represent three latent constructs, and *M* represents method variance. The *I*s (i.e., I1–I9) are indicators (e.g., scale items or scales themselves), three each to reflect the three constructs of interest. The straight arrows represent the effects of constructs and method on the items, with the paths from *M* to the items representing method variance. The curved arrows among the three latent constructs represent potential relationships among the underlying variables of interest, estimation of which is the usual goal of an investigation. The fundamental CMV issue is uncertainty in the interpretation of relationships among measured variables, as method variance contaminates measurement, typically to an unknown extent.

Figure 2
Multiple Sources of Method Variance Affecting Measured Variables



In other words, are correlations among observed variables due to relationships among the latent constructs of interest, or does CMV contribute to those correlations in part or in whole? CMV, or the shared sources of contamination as indicated by the nine paths from *M* to the items in Figure 1, has the potential to inflate observed relationships among measured variables beyond what is due to the constructs of interest, thus presenting an alternative explanation for observed relationships among measures that poses a threat to the validity of research conclusions. If this model is in fact correct, then the omission of *M* would result in the testing of a misspecified model.

Figure 1 illustrates a single source of method variance that is shared by all measures using the same method (i.e., CMV), although it does not presume that the method-to-indicator paths are necessarily equal. This is the underlying model that is generally assumed when researchers attempt to detect and control CMV, and it serves as the basis for typical criticisms of cross-sectional self-report studies. We suggest that it is unlikely all measures in an investigation would share a single, common source (or sources) of method variance. Rather, each measured variable will be influenced by a number of factors, some of which might be shared and some of which might not be. Given an investigation with several variables, it is quite likely that each would be subject to influence from a subset of all the total factors that influence the full set of variables in the study. Such a case is illustrated in Figure 2, in which there are three sources of method variance (M1, M2, and M3), each of which has a unique pattern of relationships with the items of three measured variables. From the perspective of the observed variables, each is affected by its own set of method sources. Thus the items reflecting C1 are all affected by M1, the items reflecting C2 are all affected by M1 and M2, and the items reflecting C3 are all affected by M2 and M3. In this scenario, there is a mix of common (M1 and M2) and uncommon (M3) method variance.

That variables can differ in sources of method variance can be seen by examining meta-analyses reporting results relative to variables generally considered to bias measurement. For example, Moorman and Podsakoff (1992) meta-analyzed studies linking social desirability to organizational variables. They found that about half of their variables were not significantly related to social desirability, and thus their intercorrelations would not be inflated by this method variance source. The relevance of social desirability as a source of CMV likely depends on the focal constructs under consideration. Although social desirability might inflate relations among self-references of attitudes and evaluations, it may have weaker impact on attitudes and evaluations of nonself, neutral targets (R. Johnson, Rosen, & Djurdjevic, 2011).

Although most researchers and reviewers typically focus on the potential inflation of correlations due to CMV, it is important to also consider the attenuating effects of UMV. Specifically, extraneous influences on a measured variable that are not shared act similarly to error variance and can attenuate observed correlations (Williams & Brown, 1994). When UMV is present and relationships are thus attenuated, effect sizes will be underestimated. Moreover, UMV can have significant distorting effects on complex statistics (e.g., multiple regression, structural equation modeling), and erroneous conclusions that related constructs are unrelated can occur. In addition, because UMV is systematic error, it cannot be detected as unreliability, as responses might be consistent across items within a measure and across time.

An example of UMV is supervisor ratings of organizational citizenship behavior that may be influenced by halo error (Carpenter, Berry, & Houston, 2014), as opposed to self-ratings of job satisfaction that may be influenced by the employee's transitory mood. In this scenario, method variance systematically impacts how individual raters (e.g., self vs. supervisor) respond to survey items, yet it affects each rating source and measure independently. It may, therefore, attenuate rather than inflate observed relationships between supervisor ratings of OCB and subordinate ratings of job satisfaction. Thus, although it is often assumed that separate rating sources reduce type I error associated with CMV, different rating sources may simultaneously introduce UMV, which has the potential to increase the incidence of type II errors.

How CMV and UMV Affect Correlations

With regard to how different sources of method variance might affect observed relationships between constructs, we illustrate algebraically the effects of shared (CMV) and unshared (UMV) method variance on observed zero-order correlations (for a similar argument involving control variables, see Spector & Brannick, 2011). We begin by considering two variables with a single source of method variance (X and Y correspond to two Cs in Figure 1). The formula for correlation in terms of covariances and variances is

$$r_{XY} = \frac{Cov_{XY}}{\sqrt{Var_X Var_Y}}. \tag{1}$$

We can build upon classical test theory to decompose the variance into construct, method, and error components. This theory suggests that variance in an observed measure can be partitioned into true score and error, with the true score component representing the theoretical construct of interest (Allen & Yen, 1979). This notion can be illustrated in terms of variances as

$$V_O = V_T + V_E, \tag{2}$$

where V_O= variance of observed scores, V_T= variance of true scores, and V_E= variance of error.

The interpretation of a true score, however, is an inference based on what the true score is assumed to reflect. That the systematic component of a measure reflects one and only one construct, independent of other constructs, seems dubious. Rather, it is most likely that the systematic (nonerror) variance of an observed score can be partitioned into a number of sources rather than a single true score, as in

$$V_T = V_C + \sum V_{M_i}, \tag{3}$$

where V_C is variance in the intended construct, and M_i is an influence on the observed score of method effects other than construct score and error. V_C is a theoretical entity representing what we would like to infer about the nature of a measured variable and what it represents. The M_is represent unintended sources of systematic variance due to how the particular construct is measured. They reflect alternative explanations, in part or whole, for observed relationships between variables. As we will discuss later, there can be many potential sources of additional method variance.

Substituting the right side of Equation 3 for V_T in Equation 2 gives us

$$V_O = V_C + V_E + \sum V_{M_i}, \tag{4}$$

which is a more complex view of observed scores. Acceptance of Equation 4 recognizes that we should not equate measures with constructs (cf. Binning & Barrett, 1989), assuming that construct validity is an all-or-none proposition. Instead, the systematic variance in an observed score can reflect multiple constructs (Spector & Brannick, 2009), and even though the interpretation that a measured variable reflects a theoretical construct might have supporting evidence, there is still room for the action of unintended influences that might affect observed scores.

If an observed measure contains unintended method variance, its total variance will be larger than if the measure contains only construct and error variance. Thus the denominator in Equation 1 will be inflated. If method variance is shared between *X* and *Y*, the covariance between them will be inflated as well. Because the covariance between two linear combinations is equal to the sum of the covariances among all possible pairs of terms (Nunnally & Bernstein, 1994), we can write the covariance between *X* and *Y* in terms of construct score (intended construct), error, and method components. Since errors are assumed to be random and unrelated to true scores and methods, we simplify the equation by treating these components as equal to zero. The remaining four terms are

$$Cov_{XY} = Cov_{X_C Y_C} + Cov_{X_C Y_M} + Cov_{X_M Y_C} + Cov_{X_M Y_M}, \tag{5}$$

where the C subscripts represent intended constructs and M subscripts represent method. For simplicity we show only one source of method variance.

If we make one additional simplifying assumption that methods and traits are independent (although as noted below, sometimes this assumption is incorrect), the second and third terms to the right of the equal sign drop out, so that the inflation is equal to the covariance between the method components in X and Y. If a single source of method variance is shared, the last term should be larger than zero, and thus the observed covariance will be larger than the covariance among the theoretical constructs. Thus, inflation of the covariance in the numerator will tend to increase the size of the observed correlation.

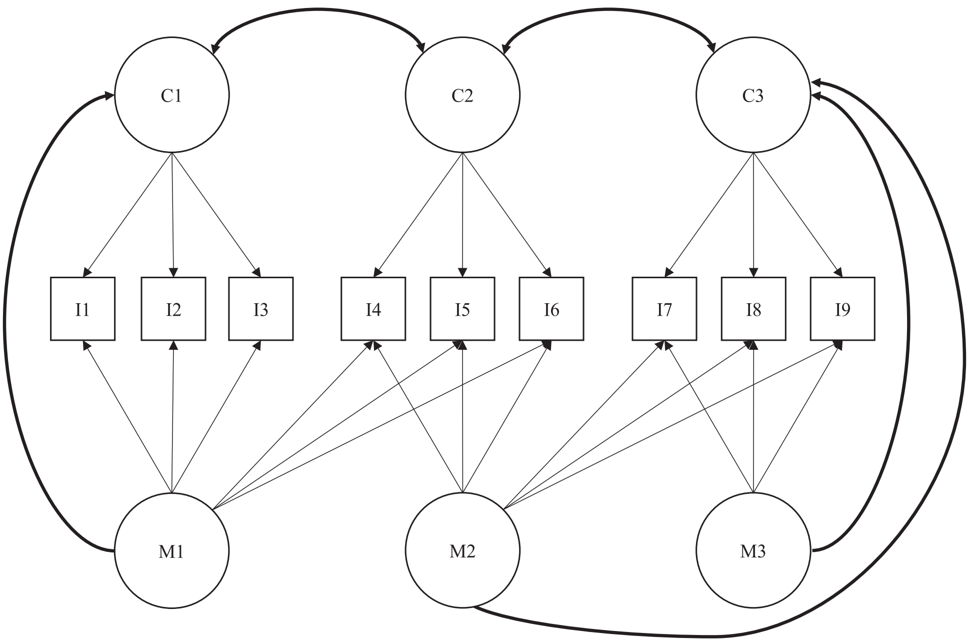
When sources of method variance are not shared, however, the covariance between X and Y is unaffected because the covariance between the X and Y method components equals zero. Instead, the variances in the denominator of Equation 1 are inflated, meaning the observed correlation will be attenuated because the denominator is larger. Furthermore, as noted earlier, there can be multiple sources of additional variance in a measure, and it is plausible that some are shared and others are not. Equation 5 could be expanded to include additional sources of shared method variance, each of which would have the potential to inflate the covariance between observed X and Y. Those sources that are not shared will increase variance in their respective observed measures, and therefore inflate the denominator of Equation 1, without also affecting the numerator. Whether the net effect is to inflate or deflate correlations depends on the relative magnitude of the shared and unshared sources, and how much the numerator and denominator are affected.

Implications of Complex Models of Method Variance

The possibility that CMV and UMV co-occur adds additional ambiguity to the interpretation of observed relationships. In some cases, one form of method variance might dominate, resulting in an over- or underestimate of construct relationships. In other cases, however, CMV and UMV might cancel one another out, leaving observed relationships as reasonably accurate reflections of underlying relationships. The need to understand the magnitude of method variance sources on observed variables is well noted by Baumgartner and Weijters (2012), who argue that different sources of method bias are likely to vary in terms of their seriousness and influence.

The incorporation of multiple sources of method variance (e.g., three in Figure 2) affecting measured variables in this way departs from conceptions derived from Campbell and Fiske (1959) in which method variance resides in methods and thus is reflected in a single latent construct for each method (i.e., as noted in Figure 1). Thus, countering conventional wisdom, we argue that there is not a single set of factors inherent in a method that produces method variance. In contrast, there exists a set of factors or sources, such as mood or social desirability (Williams & Anderson, 1994), that influences each observed measure. We refer

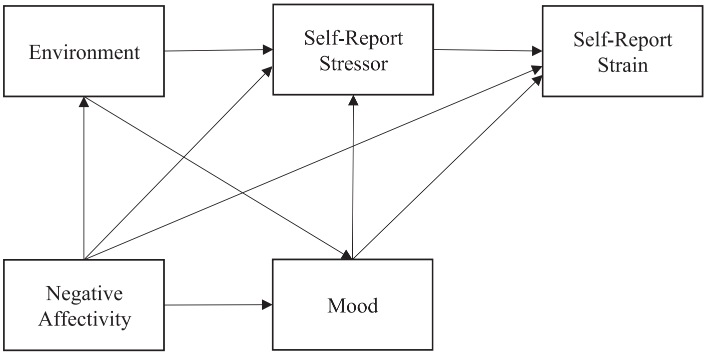
Figure 3
Multiple Sources of Method Variance Affecting Constructs of Interest and Measured Variables



to such sources as method variance to distinguish it from variance due to the intended construct, but we do not imply that method variance is purely a function of method. Rather, each measured variable can have its own unique set of extraneous influences, that is, sources of method variance.

In Figure 2 we illustrated a case in which method variance has effects only on the measurement part of the model. That is, method variance affects observed variables but not their underlying constructs. Since sources of method variance are constructs in their own right (Fiske, 1987a), relationships between substantive constructs and method sources are feasible. Figure 3 illustrates such a case in which the CMV and UMV sources have paths not only to the observed variables but to the latent constructs as well. For instance, such relationships might occur when sources of method variance are personality variables, such as negative affectivity or social desirability. In fact, social desirability was conceptualized by Crowne and Marlowe (1964) not as a measurement bias but as a personality trait of approval need. Individuals high on approval need are motivated to present themselves to others in a favorable light. Thus, they respond to scale items in a socially desirable direction. It is, however, not just that they report themselves in a favorable light; their approval need also leads them to behave in socially desirable ways, in at least their public behavior. In this case, high approval need would influence both the construct (i.e., people adjust their behavior in a socially desirable direction) and the measure (i.e., people exaggerate their ratings of socially acceptable behaviors).

Figure 4
Stressor-Strain Model Showing Complex Role of Mood and Negative Affectivity



Depending on the nature of the constructs, even more complex patterns of relationships are possible. For example, it has been shown that negative affectivity, sometimes considered a source of method variance, can contribute to the objective work environment through a variety of substantive mechanisms (Spector, Zapf, Chen, & Frese, 2000). Furthermore, Williams, Gavin, and Williams (1996) found support for models with negative affectivity playing both a contaminating (method variance) and substantive role. Add to the figure multiple sources of method variance, with connections not only with constructs but among method variance sources as well, and models can become quite complex. Such a model is illustrated in Figure 4 in the context of occupational stress. The substantive part of the model suggests that environmental stressors lead to perceptions of stressors that lead to strain. Negative affectivity, as a personality variable, and mood, as a transitory state, can both operate as sources of method variance, as signified by the paths from each one to stressors and strains. However, there is evidence that negative affectivity can be an antecedent to how employees experience the work environment (Levin & Stokes, 1989; Spector et al., 2000) and that negative affectivity can affect mood (Watson & Clark, 1984). The environment can also shape mood (Weiss & Cropanzano, 1996). Disentangling the various mechanisms and reasons for relationships would require complex approaches that are capable of isolating effects independent of method sources. In this case, methods should be employed that would be independent of the target employees' mood and personality. One might, for example, use non-self-report measures of the work environment, assuming such measures could be found that are independent of the target person's mood and negative affectivity. Of course, such measures likely contain UMV, which also requires attention.

Empirical Demonstrations

The preceding section argued that a more sophisticated view of method variance is needed, one that recognizes the influence of multiple sources of method variance, allows for these sources to impact some but not all measured variables, and recognizes that these sources may be correlated with one another, uncorrelated, or some combination of both. In this section we present two empirical demonstrations of the preceding arguments. The first involves simulations with six different examples to model the effects of varying levels of substantive variance, CMV, and UMV, as depicted in Figure 2. The second involves self-reported survey data from a sample of employees to model the potential for dual substantive and method effects on observed relationships and the challenges these dual effects pose for accurately interpreting results. The goal of these examples is to illustrate the importance of properly specifying method variance processes when testing substantive hypotheses as well as the hazards of failing to do so.

Simulation Example: The Combined Effects of CMV and UMV

For this example we utilize a common simulation technique of generating a covariance matrix from a source model, then examining the effect when we test it on a target model that is different from the source in a meaningful way (see Brannick & Spector, 1990; Richardson et al., 2009). Building on Figure 2, the simulation starts from a foundational confirmatory factor analysis (CFA) model that allows for three substantive factors (C1, C2, C3) to be correlated with each other and in which each factor is measured using three indicators (I1 through I9). In terms of method variance processes, the CFA model further specifies that three sources of method variance (M1, M2, M3) are in operation, with two of these sources influencing the indicators of two of the three substantive factors. Specifically, M1 influences the indicators of C1 and C2, while M2 influences the indicators of C2 and C3. In addition, the CFA model includes a third method factor (M3) that influences only the indicators of C3.

Across all six examples (described below), substantive factor correlations are assumed to be .20 (between C1 and C2), .30 (between C2 and C3), and .40 (between C1 and C3). For the first set of three examples, parameter values were chosen to represent conditions in which the amount of substantive variance in each indicator is assumed to be 80% of the total indicator variance, and the amount of method variance in each due to the total of method factors influencing it is 20% (Examples 1a through 1c). In the second set of three examples (2a through 2c), the amounts of substantive and total method variance are set to values of 60% and 40%, respectively. These values were chosen given their use in previous method variance simulation research (Richardson et al., 2009; Williams & Brown, 1994; Williams & O'Boyle, 2015). Within each set of examples, we further specify three distinct patterns of method factor correlations. The first pattern assumes that all three method factors are uncorrelated (Examples 1a and 2a). In Examples 1b and 2b, we specify low method factor correlations, including values of .10 (between M1 and M2), .20 (between M1 and M3), and .30 (between M2 and M3). Finally, in Examples 1c and 2c, we specify high method factor correlations, including values of .30 (between M1 and M2), .40 (between M1 and M3), and .50 (between M2 and M3). For all examples, the reliabilities of the three indicators (which represent the combined amount of substantive and method variance) for each latent variable varied and are assumed to be either .692 (for I1, I4, and I7), .500 (for I2, I5, and I8), or .368 (for I3, I6, and I9).

For Examples 1a through 1c (ratio of true to method variance = 80:20), the specified patterns of true correlations and reliabilities result in substantive factor loading values of .543, .632, and .744. For Examples 2a through 2c (ratio of true to method variance = 60:40), the factor loading values are .470, .548, and .645. When determining the method factor loadings, the percentage of method variance is assumed to be equal for all indicators of a latent variable. However, since the reliabilities of the indicators vary, so do the method factor loadings. For the C1 indicators (which are influenced only by M1), the values for the three method factor loadings linking to I1 through I3 are .271, .316, and .374 in Examples 1a through 1c, and .384, .447, and .526 for Examples 2a through 2c. The C2 and C3 indicators are each influenced by two method factor loadings. Consequently, for indicators I4 through I9, method factor loadings for the 20% method variance examples (1a through 1c) are .191, .223, and .264, while for the 40% method variance examples (2a through 2c), the values are .271, .317, and .372.

To implement the simulation, we generated the population covariance matrix for each of the six examples by using the parameter values described above as fixed values in LISREL (Jöreskog & Sörbom, 1996), and we saved the fitted covariance matrix for use in examining two models. The fitted covariance matrices for these population models were then used as input in evaluating two alternative models. The change in parameter estimates for these alternative and correctly specified models demonstrates the impact of incorrect assumptions about method variance (see Williams & O'Boyle, 2015, for another demonstration of this approach). We then used the population covariance matrix for each example to test alternative models that failed to include the multiple sources of method variance. For each example, we first examined a model that specified the substantive variables and their indicators but allowed for no method factors (referred to below as the "no-method-factor" model). The comparison of the three substantive factor correlations from this model to the original values used to generate the population covariance matrix (.20, .30, .40) demonstrates bias associated with a researcher ignoring the effects of complicated method variance processes. In addition, for each of the six examples, we evaluated a model incorporating a single method factor linked to all substantive indicators. This model is referred to as the unmeasured latent method construct model (ULMC; Richardson et al., 2009). We include it because it is equivalent to the traditional CMV model shown in Figure 1 and because it is frequently used by organizational researchers (Williams & McGonagle, 2016). The comparison of the three substantive factor correlation values in the ULMC model to the true values of .20, .30, or .40 demonstrates bias due to attempting to control for method variance that is actually multidimensional but assuming and modeling only unidimensional method variance.

Table 1 compares the true correlations on which the population covariance matrices are based with the substantive correlations as obtained from the no-method-factor and ULMC models across all six examples (i.e., three in which the ratio of true variance to method variance is 80:20 and correlations among method factors are null, lower [.10, .20, .30], or higher [.30, .40, .50]; three in which the ratio of true variance to method variance is 60:40 and the correlations among method factors are null, lower, or higher). We turn first to the results from the no-method-factor models. As described above, the observed relationship between C1 and C2 (true $r = .20$) is simultaneously inflated due to M1 and deflated due to M2 and M3. Tables 1 and 2 illustrate that the observed correlation between C1 and C2 is greater than the true value of .20 even when M1 and M2 share no variance (i.e., when $r = .00$ for the relationship between M1 and M2). As the shared variance between M1 and M2 increases and as the ratio of true variance to method variance decreases, the magnitude of the inflation becomes larger (i.e., from .30 to .33 in Examples 1a to 1c and from .40 to .46 in Examples 2a to 2c). A similar pattern of results is obtained for the observed correlation between C2 and C3 (true $r = .30$), as this relationship manifests inflation due to shared variance from M2 and attenuation due to unshared variance from M1 and M3. In contrast, because C1 and C3 (true $r = .40$) are influenced by unique method factors, the observed correlation between these factors is

Table 1
Comparison of Correlations Across Six Simulated Examples: Ratio of True Variance to Method Variance = 80:20

	True Model		No-Method-Factor Model		ULMC Model	
	C1	C2	C1	C2	C1	C2
Example 1a: No method correlations (.00, .00, .00)						
C2	.20		.30		.31	
C3	.40	.30	.32	.34	.31	.34
Example 1b: Lower method correlations (.10, .20, .30)						
C2	.20		.31		.26	
C3	.40	.30	.35	.38	.33	.33
Example 1c: Higher method correlations (.30, .40, .50)						
C2	.20		.33		.27	
C3	.40	.30	.40	.43	.37	.36

Note: ULMC = unmeasured latent method construct.

always smaller than or equal to the true correlation of .40. As would be expected, though, the magnitude of attenuation decreases in the higher method correlation conditions (i.e., r goes from .32 to .40 in Examples 1a to 1c and from .24 to .40 in Examples 2a to 2c). This pattern occurs because both C1 and C3 are also influenced by the action of C2, which produces increasing inflation as the relationships among the method factors grow stronger.

As mentioned, the results from the ULMC models indicate the effects of incorrectly modeling a single method factor when in fact there are multiple method factors influencing substantive variables, and both CMV and UMV are in operation. Note that when the method factors are uncorrelated in the population, specifying a single method factor actually further inflates the observed relationship between C1 and C2 (i.e., $r = .31$ and .41, respectively, in Examples 1a and 2a). At best, modeling a single method factor slightly reduces inflation in the observed C1-C2 correlations (i.e., $r = .26$ and .41, respectively, in Examples 1b and 2b) and in the observed C2-C3 correlations (i.e., $r = .36$ and .47, respectively, in Examples 1c and 2c) when method factors are correlated in the population. For the relationship between C1 and C3, specifying a single method factor serves to further attenuate observed correlations across all six examples (i.e., $r = .31$–.37 in Examples 1a to 1c; $r = .23$–.35 in Examples 2a to 2c). As such, the simulation results indicate that simplistically modeling one method factor in circumstances where method variance actually stems from multiple method factors with (a) distinct patterns of relationships with one another and (b) across substantive items can, in some cases, produce more erroneous results than does modeling no method factors at all.

Single-Source Data Example: Test of Stressor-Strain Model

To further illustrate the complexity of method variance sources as they relate to measures and constructs, we present a selection of structural equation modeling tests using an archival data set (Spector & Jex, 1991) that included measures of stressors (workload and interpersonal conflict), strain (physical symptoms), negative affectivity, and mood (operationalized

Table 2
Comparison of Correlations Across Six Simulated Examples: Ratio of True Variance to Method Variance = 60:40

	True Model		No-Method-Factor Model		ULMC Model	
	C1	C2	C1	C2	C1	C2
Example 2a: No method correlations (.00, .00, .00)						
C2	.20		.40		.41	
C3	.40	.30	.24	.38	.23	.38
Example 2b: Lower method correlations (.10, .20, .30)						
C2	.20		.42		.41	
C3	.40	.30	.31	.46	.29	.44
Example 2c: Higher method correlations (.30, .40, .50)						
C2	.20		.46		.41	
C3	.40	.30	.40	.53	.35	.47

Note: ULMC = unmeasured latent method construct.

as state anxiety over the prior 30 days). Surveys were completed by 232 individuals representing a wide range of job types. Although this was a single-source self-report study, for purposes of illustration we consider the workload measure as an indicator of the environment because it asks about something that is relatively factual and has been shown to yield reasonably high intersource correlations between employees and supervisors (e.g., $r = .49$; Spector, Dwyer, & Jex, 1988).

We ran two models, one showing a simple mediation chain (workload to interpersonal conflict to physical symptoms) and the other including potential sources of method variance. The first mediation model is illustrative of the typical model a researcher might want to test, namely, the potential impact of a stressor on strain, as mediated by a cognitive process. The model fit is quite good, $\chi^2(33) = 48.18$, $p < .05$, $\chi^2/df = 1.46$, SRMR = .07, comparative fit index (CFI) = .98, Tucker-Lewis index (TLI) = .97, root mean square error of approximation (RMSEA) = .05), and supports the typical mediation hypothesis (see Figure 5). The second model includes paths from negative affectivity, including one structural path to the environment itself (suggesting a substantive effect) and measurement paths to the items (or parcels) for conflict and symptoms. There is a structural path from workload to mood, and measurement paths from mood to both conflict and symptoms. The fit for this model is also quite good, $\chi^2(37) = 44.14$, $p > .05$, $\chi^2/df = 1.19$, SRMR = .04, CFI = .99, TLI = .99, RMSEA = .03), but the structural paths fail to support substantive mediation (see Figure 6). In this case, there is evidence that the apparent mediation is due to method variance, and once controlled, support disappears.

Although we illustrate mood and negative affectivity as potential sources of method variance, in this particular case it is not certain what their roles really are. For example, mood might simply be biasing the reports of conflict, but it also might serve a substantive role as either an antecedent of the conflict itself (people in negative moods are irritable and prone to conflicts) or the result (conflicts put people in bad moods). As we argue next, before appropriate strategies for dealing with method variance can be enacted, empirical work is needed to build a case for sources of method variance that goes beyond correlations.

Figure 5
Results of Mediation-Only Model

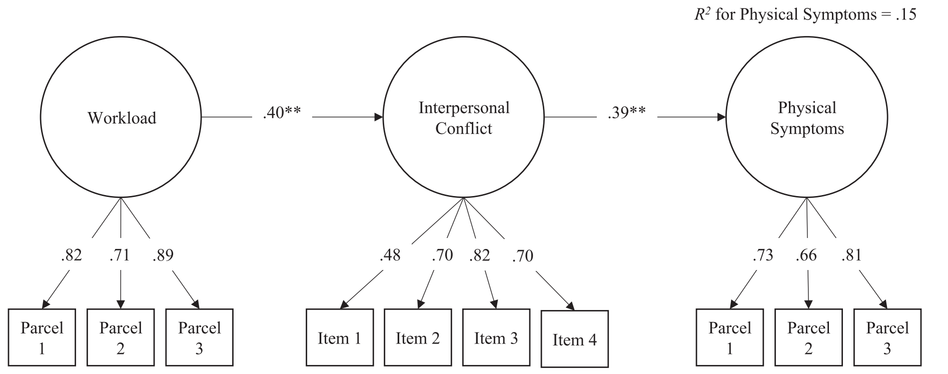
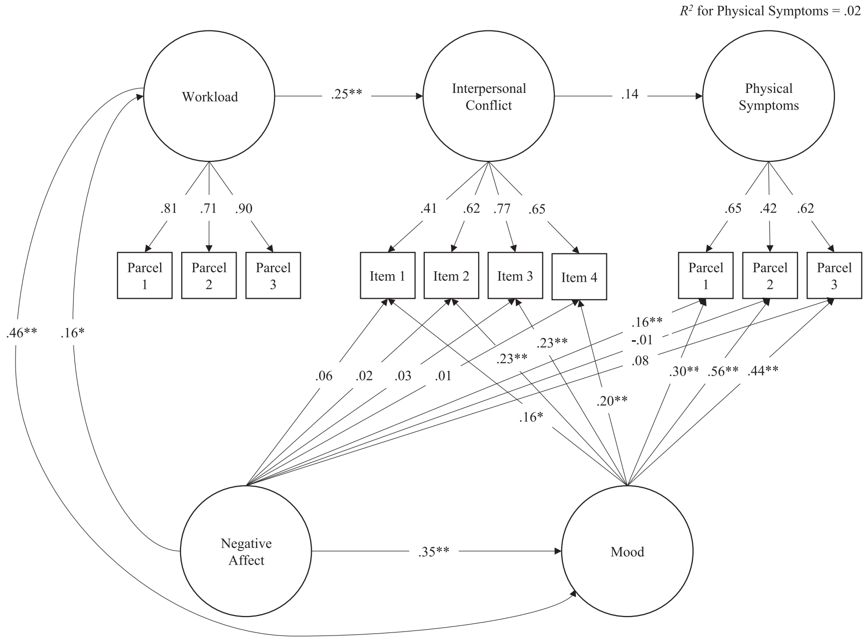


Figure 6
Results of Mediation Model With Potential Method Variance Sources Added



Identifying Method Variance: Theories of Measures

The complex nature of method variance requires a more comprehensive strategy for dealing with it than what is generally applied. Extant approaches have mostly been based on the assumption that method variance operates at the level of the method. For instance, a common device is to control for method variance by varying the methods by which variables in a model are measured or obtained. Our new conception of method variance as operating at the level of the measured variable suggests a different tactic. To fully interpret the meaning of a measure and its relationships with other variables, we must first understand the sources of variance that affect it. The field needs to build a knowledge base upon which to rely in deciding strategies for controlling method variance in investigations. Although there have been calls to consider method variance at the outset of a study and measure design (Baumgartner & Weijters, 2012; MacKenzie & Podsakoff, 2012), knowledge about what sources to control based on the constructs and methods under consideration is incomplete.

Effectively identifying and controlling CMV and UMV in our studies requires that we develop theories for each measure that enable us to specify sources of variance. Such an approach was suggested in a more general manner by Schmitt (1994), whose Table 1 illustrated various sources of method variance for categories of variables (e.g., job attitudes or personality). We go a step further and argue that each individual measure can have its own set of method variance sources, such that we cannot assume that all personality variables or all behavioral checklists share the same sources. For example, social desirability, which is often positioned as an important source of method variance, does not relate consistently across organizational variables (Moorman & Podsakoff, 1992). Within the personality domain, social desirability correlates with measures of emotional traits, such as negative and positive affectivity (Soubelet & Salthouse, 2011), but not necessarily with cognitive traits, such as locus of control (Spector, 1988). Furthermore, even within instruments and scales, social desirability might affect some items more than others (Chen, Dai, Spector, & Jex, 1997).

Theories of measures have to consider the context in which measurement occurs, that is, the broader method in which it is embedded. We can consider a method to assess a construct as consisting of everything an investigator does to quantify it (Fiske, 1987b). For example, Fiske (1982) noted that the method for a self-report scale includes not only the format of the scale but also characteristics of the examiner and setting, instructions, and the expressed reason for the assessment. Doty and Glick (1998) divided method characteristics into three categories: measurement techniques (e.g., response formats, scale anchors, communication medium), data sources (e.g., self or supervisor), and time frame (e.g., past week or past year). Each characteristic of the method used to assess something in combination with the content of what is being assessed has the potential to be influenced by one or more extraneous sources of unintended variance.

A Strategy for Building a Measure Theory

Building a theory of a particular measure that specifies potential sources of method variance and the conditions under which such sources will manifest is a time-consuming empirical process that requires the integration of different forms of evidence. An example of combining a diverse set of findings to build a case for substantive relationships for a class of variables (i.e., job conditions) based on existing findings was provided by Spector (1992). A similar approach could be used for individual measures, combining evidence from the literature with additional studies conducted for the purpose of investigating method variance. We suggest a five-step procedure that includes both literature review and primary studies. These steps are not remedies that can be applied in a single study to control the effects of method variance but rather serve to identify sources of method variance that can inform how best to control it in subsequent studies. It involves considering the nature of the construct and measure, including the context of measurement. Both primary and secondary research is required, integrating existing knowledge with results of new studies. The goal is to catalog sources of method variance and conditions under which they will manifest that will inform research design choices, as well as selection of measures to include as method variance controls.

Step 1: Literature review to identify potential sources. The first step in building a measure theory is to consult the literature for suggestions about the kinds of method variance sources that might be relevant to a particular measure or type of measure. As noted previously, one must not assume that correlations with variables suspected of being method variance sources automatically indicate method variance. At this point the goal is to identify potential sources of method variance to be addressed in later steps.

There are many features of a measure that can affect the sources of method variance that are relevant to consider. The majority of measures in organizational research rely on humans as instruments (i.e., human judgment is used to quantify our variables). Table 3 presents a breakdown of characteristics and features of such methods based on human judgment. Listed here are a number of factors that can help determine the sources of method variance. The table organizes them by categories of source (e.g., self vs. other), whether the item is open-ended, the nature of the response choices, the communication channel, and context. The sources of method variance are determined in part by these method features, such that their presence or absence will determine how susceptible a particular measure is to method variance sources. For example, a person's standing on a given personality trait might affect his or her reports about the self but not about others. Extraneous factors that might influence human judges would be different from those that affect computer algorithms for content analysis. This list is far from comprehensive, but its length is indicative of the complexity of method characteristics that need to be considered when specifying potential sources of method variance for each measure.

An important characteristic is the source of the rating, that is, whether people are being asked to report on themselves or on others. When others are asked to report on a target, one should consider whether the ratings are provided by an internal (e.g., coworker, subordinate, or supervisor) or external (e.g., client, customer, patient) rater. Alternatively, ratings may be provided by a member of the research team who is aware of the purpose of the research. These raters vary in terms of their training and other characteristics that may influence their ratings. Of course, organizational researchers also use physiological measures that do not rely on human judgment, whether assessed directly from individuals, such as an individual's blood pressure, cholesterol level, or weight (Ganster & Rosen, 2013), or taken from organization records, such as days absent or sales volume. Despite the absence of human judgment, such measures can still be affected by extraneous factors that are not of interest in a given study (for a discussion of biases in physiological measures, see Fried, Rowland, & Ferris, 1984), and they can be shared with self-reports. To illustrate, ingesting caffeine can serve as a CMV source between physiological measures and self-report measures. Caffeine can have an effect on heart rate and blood pressure as well as mood that might then affect self-reports of many variables. Thus relationships between physiology and self-reports might be inflated by CMV attributed to caffeine consumption. The possibility of method bias is therefore not restricted to measures that rely on human judgment.

Table 3
Characteristics and Critical Features of Methods

Method Characteristic	Critical Features
Rating source	Self vs. other
	Awareness of research purpose
	Rater characteristics
	Rater training
Open-ended response	Content analysis
	Human vs. machine coded
Rated response	Estimation of construct vs. checklist of occurrence
	Item formats
	Judgment type (agreement, evaluation, frequency)
	Number of response choices
	Bipolar vs. unipolar
	Item wording direction (unidirectional vs. bidirectional)
	Instructions
	Response time frame (e.g., past week or month)
Communicational channel	Interview
	Degree of structure
	Medium
	Face-to-face
	Phone
	Other media
	Characteristics of interviewer
	Instructions
	Setting (e.g., home vs. work)
	Questionnaire
	Medium
	Paper and pencil
	Computer in a lab
	Web based
	Location
	Device (e.g., desktop, laptop, phone, tablet)
Context	Privacy
	Anonymous vs. confidential
	Who has access to results
	Vulnerability to punishment for negative responses
	Purpose
	Internal (e.g., management) vs. external (e.g., research)
	Incentives
	Extrinsic (e.g., cash, gift card)
	Intrinsic (e.g., serving a larger purpose)
	Employee participation

The literature contains a great deal of information about likely method variance sources that can serve as a starting point to guide subsequent steps (see Table 2 in Podsakoff, MacKenzie, Lee, & Podsakoff, 2003, for an extensive list of potential sources). For self-report ratings, the focus has been largely on individual differences, most notably, personality.

Affective traits, such as negative affectivity (Watson, Pennebaker, & Folger, 1986), and cognitive traits, such as hostile attribution bias (Spector & Zhou, 2011), are likely candidates to affect ratings about the workplace. Additional sources of method variance are associated with the design of a measure. Item wording direction, for example, can affect responses to items (Alessandri, Vecchione, Donnellan, & Tisak, 2013), with items written in the same direction being more highly correlated with one another than with items written in the opposite direction (Spector, Van Katwyk, Brannick, & Chen, 1997). Furthermore, the degree of abstractness of constructs and items can be important, as abstract items require greater cognitive load and leave more room for idiosyncratic judgments (Doty & Glick, 1998).

As part of the literature review, special attention should be given to the nature of the constructs under consideration. In Table 4, we illustrate the connection between characteristics of measures according to their nature (factual, perceptual, affective/attitudinal, behavioral, or evaluative). As the table illustrates, the construct's nature influences the main sources used to assess it (e.g., self vs. other rater), which plays a critical role in determining a construct's susceptibility to different sources of method variance. Perceptual measures, for example, assess characteristics of the work environment (e.g., job characteristics or supervisory style), and they can be gauged through self-reports, reports by colleagues/supervisors, or by observers, which determines the relevance of different method variance sources. When measured by others, for example, there can be rater biases (e.g., liking of the target) that can be important. On the other hand, affect and attitudes can best be assessed with self-reports, since it is difficult for others to observe internal states of an individual. Those self-report measures might be affected by mood, certain personality traits (e.g., neuroticism), and response sets.

If the measure is well established and has been extensively used, there might be considerable data that inform that measure's theory. Searches of the literature can locate primary studies using the measure of interest that also include measures of conceivable sources of method variance. There also might be studies that address issues of bias and construct validity. If sufficient studies are available, meta-analytic methods could be used to compute mean effect size as an indicator of potential amount of method variance and to explore potential boundary conditions and moderators. Meta-analyses can also be helpful in exploring the possible impact of potential sources of method variance, such as mood, negative affectivity, or social desirability. Some meta-analyses will report results separately for specific measures if there are enough studies using them (e. g., Bowling & Beehr, 2006). In most cases there will be a variety of measures of the same or similar constructs that are combined, so the meta-analysis will not provide specific information about the measure in question. This can still be a good start, though, as meta-analytic results can indicate whether or not potential sources are related to similar measures. Furthermore, meta-analyses can provide a list of studies that could be scanned for those that used the measure of interest. Finally, in addition to mean correlations, moderator tests can provide information about boundary conditions under which method variance sources might have effects.

Step 2: Conduct primary studies of method variance sources. Where studies do not exist linking a potential source of method variance to a measure of interest, primary studies should be conducted to provide such data. Given the large number of potential sources of method variance, empirical work is certainly needed to link these sources to specific measures and to identify the magnitude of effects for each source (Baumgartner & Weijters, 2012). This step

Table 4

Connection Between Constructs and Methods

Nature of Construct	Factual	Perceptual	Affective/Attitudinal	Behavioral	Evaluative
Types	• Demographics • Information from records • Medical diagnosis • Objective work conditions • Work schedule	• Job characteristics • Job stressors • Leadership style	• Emotions • Intentions • Job Attitudes • Pain • Physical symptoms	• CWB • OCB	• Job performance
Examples	• Age • Absence • Span of control • Training record	• Charismatic style • Group cohesion • Role ambiguity • Task significance	• Angry mood • Back ache • Job satisfaction • Turnover intentions	• Reports of behavior such as absence, helping a coworker, sabotaging work, volunteering for extra work	• Ratings of performance quality
Main sources	• Self • Coworker • Supervisor • Records	• Self • Coworker • Supervisor • Observer	• Self	• Self • Coworker • Supervisor	• Self • Coworker • Supervisor • Observer
Potential sources of method variance	• Systematic recording errors • Intentional distortion	• Biases of rater: liking of target, mood, social desirability for self-ratings, response sets	• Mood • Social desirability • Response sets	• Mood • Limited information • Social desirability • Impression management • Fear of revealing negative behavior (self-ratings)	• Rater errors • Impression management (self-ratings)
Comments	• Low bias • Limited to things observable/verifiable	• Influenced by subjectivity • Better measure of perception than objective environment • Reverse effects where affect/attitudes affect perceptions of environment	• Difficult to know if real or imagined • Subject to intentional and unintentional distortions	• Behaviors vary in sensitivity and likelihood people will reveal • Other raters have limited chance to observe some behaviors • OCB subject to impression management distortion (people pretending to perform)	• Raters often can only provide general impressions • Raters can be reluctant to give honest appraisals due to political reasons

Note: CWB = counterproductive work behavior; OCB = organizational citizenship behavior.

is necessary for new measures and should be a routine part of scale development. Even for existing scales, there might not exist enough studies to cover all potential sources of interest.

Step 3: Comparison of different measures and sources. Where possible, one should compare results of a measure of interest with alternative measures that would not be expected to share one or more sources of method variance. The alternative source might be a rating by another person (e.g., coworker or supervisor), or it might be a measure that does not rely on human judgment, such as a physiological measure. Finding a correlation between the measure of interest and an alternative that does not share a source of method variance is evidence that the two measures reflect a common construct at least to some extent. Given that alternative source correlations are typically small, however, it does not allow one to totally rule out that the measure in question is free of method variance. It should also be kept in mind that measures based on very different methods (e.g., self-report of sleep quality vs. physiological measure of sleep) likely each contain UMV, thus attenuating correlations between alternative measures.

Assuming alternative measures are related, one would then compare them in their relationships with criterion measures of other constructs. Finding a similar pattern of relationships of alternative measures with criteria expected to relate would provide evidence that observed correlations are not attributable to CMV alone. Of course, this interpretation assumes that the two alternative measures do not share sources of CMV, which is something that can be difficult to conclusively rule out. Nonetheless, comparing results with the measure of interest to a variety of alternative measures can help build a case that observed relationships are not due to CMV. It should also be kept in mind that UMV can attenuate correlations, and some measures might have more UMV than others, thus yielding smaller observed relationships. Many of the constructs of interest in the organizational sciences are internal states (cognitions or emotion) or are experiences best known to target individuals. Thus reports by others are likely to be less accurate and thus yield smaller correlations.

Step 4: Distinguish method from substance. The first three steps are useful for identifying potential sources of method variance for a measure and to help indicate whether observed correlations of that measure with alternative sources have the potential to be attributed to CMV. A limitation is that the steps merely show patterns of relationships that might or might not be consistent with a method variance explanation. Additional strategies that can be accomplished through a variety of experimental (both laboratory and field) and nonexperimental methods are needed to distinguish method from substance. This avenue requires controlling and/or manipulating both the construct of interest and the source of method variance to see if the measure of interest reflects the construct and/or the source of method variance.

An experimental approach directly introduces control/manipulation to the construct and method variance source or sources. For example, suppose the target measure is a self-report of workload. A laboratory experiment could be designed to hold workload constant, with participants first working at a task for a period of time before completing the workload measure. Mood could then be explored as a potential source of method variance. In a first stage, mood might be assessed to see if it relates to perceived workload. If it does, this finding would suggest the possibility of mood impacting the workload measure, but it also could reflect the impact of workload on mood. Individuals who find the task difficult might perceive the workload to be higher, which might affect their mood. A more conclusive experiment would be to manipulate mood to see if reports of workload are affected. This could be done with mood induction manipulations before and after the task. Finally, the task itself could be manipulated by systematically varying objective workload to verify that the workload ratings are impacted by objective features of the task vis-à-vis mood. Through this series of experiments one could show the extent to which objective features of the task and mood impact the workload report. Such experiments could be used for a variety of constructs that can be experimentally manipulated in both laboratory and field settings.

Nonexperimental methods can also be used to address method variance in measures. Although control is more limited compared to experimental studies, the idea is to investigate responses to a measure as the substantive variable and potential method sources are controlled or vary systematically. This investigation might be accomplished through structured observations over time. Considering the workload-mood example, one might use a daily diary study to explore workload reports as a function of objective workload changes and mood. In jobs with a constant workload (e.g., a machine-paced assembler job), one could see if mood fluctuations from day to day predict workload reports. For a job with a varying workload (e.g., sales clerk with different numbers of customers), the effect of objective workload on workload reports could be shown. Conducting such studies with a variety of samples that differ in workload variance could be used to shed light on the extent to which ratings correspond to actual workload and mood operates as a source of CMV.

Even with variables that do not allow for objective assessment, it is possible to use daily diary and other time-based designs to investigate CMV in measures. The daily fluctuation of variables could be modeled by potential sources of CMV. For example, job satisfaction daily fluctuation could be related to mood. Of course, as with earlier steps, merely finding that mood relates to job satisfaction does not necessarily rule out alternative possibilities (e.g., that on days when something distressing happens at work, both job satisfaction and mood are adversely affected). However, this approach can contribute to building a theory of a measure and what might affect it.

Step 5: Response styles and systematic rater error. Response styles (i.e., the tendency to respond to survey items in particular ways regardless of item content) are also potential sources of method variance that differ across individuals. Baumgartner and Steenkamp (2001) noted five forms of response styles for Likert-style agreement scales: acquiescence (tendency to agree with items), disacquiescence (tendency to disagree with items), extreme responding (tendency to choose the highest or lowest rating), midpoint responding (tendency to choose middle ratings), and noncontingent (careless or random responding; see Meade & Craig, 2012). Baumgartner and Steenkamp note methods to assess these styles that can be used to help identify if particular measures might be subject to this potential source of CMV.

For evaluative reports (e.g., ratings of job performance), rater errors have been noted. Similar to response styles, rater errors represent patterns of response across different items regardless of content. Between-target rating errors are central tendency (choosing middle ratings), leniency (choosing favorable ratings), and severity (choosing unfavorable ratings). Within-target rating error is halo, that is, the tendency to rate the same across all dimensions. Although studied mainly with other-reports, these rater errors can also apply to self-ratings. It should be noted that while rating errors (e.g., halo) have been discussed as method variance sources (i.e., biases in ratings), they can reflect relationships among constructs, or what has been referred to as true halo (Murphy, Jako, & Anhalt, 1993).

Response styles and rater errors do not differ only among individuals, but there are country and cultural differences in them as well (Spector, Liu, & Sanchez, 2015). For example, people in collectivist societies tend to exhibit more acquiescence (T. Johnson, Kulesa, Cho, & Shavitt, 2005). Asians, such as Chinese, tend to make more modest ratings of self-performance than do Americans (Farh, Dobbins, & Cheng, 1991). Thus, if a sample includes individuals from different cultural or national backgrounds, culture might serve as a source of method variance that needs to be considered.

Once response styles are shown to be a potential source of method variance, the effects on observed relationships need to be identified. Even when measures are subject to response styles, those styles do not necessarily carry over from one measure to another (Rorer, 1965). An experimental strategy might also prove useful in exploring response styles as a source of CMV. The response formats of scales (e.g., agreement vs. frequency, summated rating vs. forced choice) can be systematically varied in order to see if format affects the relationships of a measure with other measures. Such an approach was used by Spector, Bauer, and Fox (2010) to show that the observed correlation between counterproductive work behavior and organizational citizenship behavior might have been affected by rating format.

Strategies for Controlling Identified Method Variance

Knowledge about sources of method variance should be incorporated into the design of investigations so that potential biasing and contaminating factors can be controlled. This can be done in the choice of measures and procedures that introduce control in the design or in the inclusion of method variance measures to allow statistical control. When feasible, design control is generally preferred to eliminate method variance sources (for an extensive list of potential design or "procedural" strategies, see Podsakoff, MacKenzie, & Podsakoff, 2012). It should be kept in mind, however, that design strategies need to be based on the nature of the measures and constructs (i.e., the measure theories) for a particular investigation. Furthermore, there can be trade-offs in the choice of strategies that suggest the need for multiple approaches and multiple studies to provide convincing evidence for particular conclusions. To wit, sources of method variance that are idiosyncratic to individuals might be addressed through the use of different raters (e.g., employee and supervisor). Such choices, however, should not be based on the assumption that the use of different raters is a total solution, as it controls for only some sources of method variance but not others. Furthermore, the use of different raters can represent a trade-off between CMV that might occur within source and UMV produced by using different sources. Nevertheless, design strategies can be very useful in dealing with method variance.

It is not often possible to control all sources of method variance through design approaches alone, necessitating the use of a statistical control approach. Statistical control requires including one or more measures to capture specific, hypothesized sources of method variance, such as a measure of social desirability or inclusion of a set of neutral items (e.g., Judge & Bretz's [1993] neutral objects scale) that can be used to identify and control for extreme responding. The assessment of multiple method variance sources allows for the testing of models that are more complex than the traditional one, where different sources might have different effects across measured variables at either the scale or item level (Simmering, Fuller, Richardson, Ocal, & Atinc, 2015). When specifying such models, relationships between a method variance source and substantive constructs can be determined.

Note that control should be incorporated for both shared and unshared sources of method variance to acquire more accurate estimates of construct relationships. With predictor models, such as multiple regression, an unshared source of method variance that is entered into an analysis can be considered a suppressor variable that relates to another predictor but not the criterion. The included suppressor variable controls UMV and thus yields a more accurate estimate of construct relationships. It should be kept in mind, though, that one should include variables as potential suppressors because there is evidence that they serve as a source of method variance. It should not be assumed if a variable is related to the predictor but not the criterion that it is inherently a source of method variance. We need additional evidence beyond a pattern of relationships in order to draw that conclusion.

An example of the statistical approach was provided by Williams and Anderson (1994), who conducted comparative model tests including and omitting measures of negative and positive affectivity (see Hoge & Bussing, 2004, for a similar example with negative affectivity and sense of coherence). Both studies show that the sources of method variance do correlate with the substantive variables in the study, but their inclusion has only minimal effects on the structural models of interest. The presence of identified sources of method variance gives some confidence that the path estimates in structural models are not affected by method variance.

Statistical control is not an automatic process that necessarily leads to correct inference (Spector & Brannick, 2011). To properly specify a tested model, one must be confident about the role each variable plays. As in our earlier example, we might assume that variables like mood and negative affectivity necessarily play only a method variance role, but they almost certainly play a substantive role as well. Thus caution is warranted regarding how statistical control is used and interpreted.

Conclusion

We argue here that the most productive way of dealing with method variance is to approach it not from the broad perspective of the method but from the more microscopic vantage point of the individual measure. Strategies focused on the method of measurement have led to remedies that are generic and fail to fully consider that method variance works in a much more measure-specific manner than is generally allowed. As our simulation shows, controlling for single-source method variance can in some cases result in less accurate estimates of construct intercorrelations than testing models with no controls at all. Furthermore, as we (and others in the past) have demonstrated, method variance can attenuate as well as inflate observed correlations depending on whether it is common (shared across measures) or uncommon (unique across measures). To more conclusively address method variance, one must work at the level of the measured variable to first identify sources of method variance and then to devise informed strategies to control it.

Although there is a great deal of knowledge in the literature about potential sources of method variance and how they might affect measurement, much of that literature is focused on classes of measures (e.g., personality measures) rather than specific measures. Even where individual measures are investigated, the literature on them tends to be diffuse, and it does not focus on generating a theory of the particular measure that specifies sources of variation in addition to the target construct. Papers that report on scale development are more concerned with evidence that supports construct validity than with evidence for additional sources of variance. Efforts to thoroughly investigate specific measures are needed to shed light on sources of method variance.

Building a theory of a measure is not an easy task, as it requires a literature review to compile existing knowledge as well as a series of primary studies to fill in gaps in that knowledge. Likely, the thorough investigation of a particular measure will be done by multiple research groups that independently contribute to the development of the theory. It is plausible that competing measures of the same or similar constructs using the same or similar methods will prove to have the same method variance sources. As such, it might not be necessary to conduct extensive research on every individual measure ever devised or that will ever be devised. However, we cannot rely on blanket investigations of broad classes of variables, like personality or stressors, as there is certainly variability in sources of method variance across measures within the same class. As we begin to build measure theories, the extent to which measures share theories and the characteristics of measures that affect theories will become clear. In the end, perhaps we can develop a meta-theory of measures that can specify sources of method variance depending on characteristics of constructs and how they are assessed.

When the issue of method variance is raised, the underlying question is "Why do variables correlate?" The field has come up with a variety of potential remedies (MacKenzie & Podsakoff, 2012; Podsakoff et al., 2012) that, in the final analysis, are not sufficient to fully address this complex issue. Most such remedies are generic and based on the presumption that method variance is a function of the method, and they focus primarily on CMV while overlooking UMV. In many cases, such remedies lead to a false sense of security when they show that a given relationship does not appear to be the product of CMV. Better solutions are to focus at the level of the measured variable and to consider both CMV and UMV. A control strategy that targets identified sources of extraneous variance will lead to far more conclusive answers about why variables correlate than has generally been offered in the past.

References

Alessandri, G., Vecchione, M., Donnellan, B. M., & Tisak, J. 2013. An application of the LC-LSTM framework to the self-esteem instability case. *Psychometrika*, 78(4): 769-792.

Allen, M. J., & Yen, W. M. 1979. *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.

Baumgartner, H., & Steenkamp, J.-B. E. 2001. Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38(2): 143-156.

Baumgartner, H., & Weijters, B. 2012. Commentary on "Common Method Bias in Marketing: Causes, Mechanisms, and Procedural Remedies." *Journal of Retailing*, 88(4): 563-566.

Binning, J. F., & Barrett, G. V. 1989. Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology*, 74(3): 478-494.

Bowling, N. A., & Beehr, T. A. 2006. Workplace harassment from the victim's perspective: A theoretical model and meta-analysis. *Journal of Applied Psychology*, 91(5): 998-1012.

Brannick, M. T., & Spector, P. E. 1990. Estimation problems in the block-diagonal model of the multitrait-multimethod matrix. *Applied Psychological Measurement*, 14(4): 325-339.

Campbell, D. T., & Fiske, D. W. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2): 81-105.

Carpenter, N. C., Berry, C. M., & Houston, L. 2014. A meta-analytic comparison of self-reported and other-reported organizational citizenship behavior. *Journal of Organizational Behavior*, 35(4): 547-574.

Chen, P. Y., Dai, T., Spector, P. E., & Jex, S. M. 1997. Relation between negative affectivity and positive affectivity: Effects of judged desirability of scale items and respondents' social desirability. *Journal of Personality Assessment*, 69(1): 183-198.

Crowne, D. P., & Marlowe, D. 1964. *The approval motive*. New York: John Wiley.

Doty, D., & Glick, W. H. 1998. Common methods bias: Does common methods variance really bias results? *Organizational Research Methods*, 1(4): 374-406.

Farh, J.-l., Dobbins, G. H., & Cheng, B.-S. 1991. Cultural relativity in action: A comparison of self-rating made by Chinese and U.S. workers. *Personnel Psychology*, 44(1): 129-147.

Fiske, D. W. 1982. Convergent-discriminant validation in measurements and research strategies. *New Directions for Methodology of Social & Behavioral Science*, 12: 77-92.

Fiske, D. W. 1987a. Construct invalidity comes from method effects. *Educational and Psychological Measurement*, 47(2): 285-307.

Fiske, D. W. 1987b. On understanding our methods and their effects. *Diagnostica*, 33(3): 188-194.

Fried, Y., Rowland, K. M., & Ferris, G. R. 1984. The physiological measurement of work stress: A critique. *Personnel Psychology*, 37(4): 583-615.

Ganster, D. C., & Rosen, C. C. 2013. Work stress and employee health: A multidisciplinary review. *Journal of Management*, 39(5): 1085-1122.

Hoge, T., & Bussing, A. 2004. The impact of sense of coherence and negative affectivity on the work stressor-strain relationship. *Journal of Occupational Health Psychology*, 9(3): 195-205.

Johnson, R. E., Rosen, C. C., & Djurdjevic, E. 2011. Assessing the impact of common method variance on higher order multidimensional constructs. *Journal of Applied Psychology*, 96(4): 744-761.

Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. 2005. The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology*, 36(2): 264-277.

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

Judge, T. A., & Bretz, R. D. 1993. Report on an alternative measure of affective disposition. *Educational and Psychological Measurement*, 53(4): 1095-1104.

Levin, I., & Stokes, J. P. 1989. Dispositional approach to job satisfaction: Role of negative affectivity. *Journal of Applied Psychology*, 74(5): 752-758.

MacKenzie, S. B., & Podsakoff, P. M. 2012. Common method bias in marketing: Causes, mechanisms, and procedural remedies. *Journal of Retailing*, 88(4): 542-555.

Meade, A. W., & Craig, S. B. 2012. Identifying careless responses in survey data. *Psychological Methods*, 17(3): 437-455.

Moorman, R. H., & Podsakoff, P. M. 1992. A meta-analytic review and empirical test of the potential confounding effects of social desirability response sets in organizational behaviour research. *Journal of Occupational and Organizational Psychology*, 65(2): 131-149.

Murphy, K. R., Jako, R. A., & Anhalt, R. L. 1993. Nature and consequences of halo error: A critical analysis. *Journal of Applied Psychology*, 78(2): 218-225.

Nunnally, J. C., & Bernstein, I. H. 1994. *Psychometric theory*. New York: McGraw Hill.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5): 879-903.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. 2012. Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63: 539-569.

Richardson, H. A., Simmering, M. J., & Sturman, M. C. 2009. A tale of three perspectives: Examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods*, 12(4): 762-800.

Rorer, L. G. 1965. The great response-style myth. *Psychological Bulletin*, 63(3): 129-156.

Schmitt, N. 1994. Method bias: The importance of theory and measurement. *Journal of Organizational Behavior*, 15(5): 393-398.

Simmering, M. J., Fuller, C. M., Richardson, H. A., Ocal, Y., & Atinc, G. M. 2015. Marker variable choice, reporting, and interpretation in the detection of common method variance: A review and demonstration. *Organizational Research Methods*, 18(3): 473-511.

Soubelet, A., & Salthouse, T. A. 2011. Influence of social desirability on age differences in self-reports of mood and personality. *Journal of Personality*, 79(4): 741-762.

Spector, P. E. 1988. Development of the Work Locus of Control Scale. *Journal of Occupational Psychology*, 61(4): 335-340.

Spector, P. E. 1992. A consideration of the validity and meaning of self-report measures of job conditions. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology 1992:* 123-151. West Sussex, UK: Wiley.

Spector, P. E., Bauer, J. A., & Fox, S. 2010. Measurement artifacts in the assessment of counterproductive work behavior and organizational citizenship behavior: Do we know what we think we know? *Journal of Applied Psychology*, 95(4): 781-790.

Spector, P. E., & Brannick, M. T. 1995. The nature and effects of method variance in organizational research. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology 1995*: 249-274. West Sussex, UK: Wiley.

Spector, P. E., & Brannick, M. T. 2009. Common method variance or measurement bias? The problem and possible solutions. In D. Buchanan & A. Bryman (Eds.), *The Sage handbook of organizational research methods*: 346-362. Thousand Oaks, CA: Sage.

Spector, P. E., & Brannick, M. T. 2011. Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods*, 14(2): 287-305.

Spector, P. E., Dwyer, D. J., & Jex, S. M. 1988. Relation of job stressors to affective, health, and performance outcomes: A comparison of multiple data sources. *Journal of Applied Psychology*, 73(1): 11-19.

Spector, P. E., & Jex, S. M. 1991. Relations of job characteristics from multiple data sources with employee affect, absence, turnover intentions, and health. *Journal of Applied Psychology*, 76(1): 46-53.

Spector, P. E., Liu, C., & Sanchez, J. I. 2015. Methodological and substantive issues in conducting multinational and cross-cultural research. *Annual Review of Organizational Psychology and Organizational Behavior*, 2(1): 101-131.

Spector, P. E., Van Katwyk, P. T., Brannick, M. T., & Chen, P. Y. 1997. When two factors don't reflect two constructs: How item characteristics can produce artifactual factors. *Journal of Management*, 23(5): 659-677.

Spector, P. E., Zapf, D., Chen, P. Y., & Frese, M. 2000. Why negative affectivity should not be controlled in job stress research: Don't throw out the baby with the bath water. *Journal of Organizational Behavior*, 21(1): 79-95.

Spector, P. E., & Zhou, Z. E. 2011. *Individual differences in counterproductive work behavior: The interaction of gender and personality*. Tampa: University of South Florida.

Watson, D., & Clark, L. A. 1984. Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, 96(3): 465-490.

Watson, D., Pennebaker, J. W., & Folger, R. 1986. Beyond negative affectivity: Measuring stress and satisfaction in the workplace. *Journal of Organizational Behavior Management*, 8(2): 141-157.

Weiss, H. M., & Cropanzano, R. 1996. Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior*: 1-74. Greenwich, CT: JAI Press.

Williams, L. J., & Anderson, S. E. 1994. An alternative approach to method effects by using latent-variable models: Applications in organizational behavior research. *Journal of Applied Psychology*, 79(3): 323-331.

Williams, L. J., & Brown, B. K. 1994. Method variance in organizational behavior and human resources research: Effects on correlations, path coefficients, and hypothesis testing. *Organizational Behavior and Human Decision Processes*, 57(2): 185-209.

Williams, L. J., Gavin, M. B., & Williams, M. L. 1996. Measurement and nonmeasurement processes with negative affectivity and employee attitudes. *Journal of Applied Psychology*, 81(1): 88-101.

Williams, L. J., & McGonagle, A. K. 2016. Four research designs and a comprehensive analysis strategy for investigating common method variance with self-report measures using latent variables. *Journal of Business and Psychology*, 31(3): 339-359.

Williams, L. J., & O'Boyle, E. H. 2015. Ideal, nonideal, and no-marker variables: The confirmatory factor analysis (CFA) marker technique works when it matters. *Journal of Applied Psychology*, 100(5): 1579-1602.