

# **WILL THE REAL CORRELATION PLEASE STAND UP? AN EXAMINATION OF THE EFFECTIVENESS OF STATISTICAL CORRECTIONS FOR COMMON METHOD VARIANCE USING DATA SIMULATION**

Hettie A. Richardson, Louisiana State University (hricha4@lsu.edu)

Marcia J. Simmering, Louisiana Tech University

Michael C. Sturman, Cornell University

## **ABSTRACT**

A concern of researchers is the risk of measurement error due to common method variance (CMV) when using self-reported data. The present study addresses this concern by empirically comparing four techniques for correcting CMV. Eighteen simulated datasets, with varying degrees of method variance, group agreement, and reliability, were analyzed. Based on these analyses, benefits and drawbacks of correcting CMV using the different techniques are detailed. Recommendations for using the different techniques are also provided.

In recent years, several post hoc statistical techniques have been proposed as means of identifying and correcting for common method variance (CMV) in same-source, self-reported data. Few researchers, however, appear to actually use many of these techniques to test for and deal with potential CMV problems. We believe that a key reason for researchers' reluctance to embrace these methods is unresolved concerns about the utility of these approaches for addressing method bias—concerns which can only be addressed empirically. Perhaps the most important of these concerns is that researchers cannot necessarily know the extent to which the post-hoc corrections remove true or other variance in addition to method variance (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; Spector, Zapf, Chen, & Frese, 2000). The purpose of the present study is to address these issues by using simulated data to compare four techniques for correcting bias due to CMV. The four techniques we consider are: partial correlation (in which CMV, as represented by a measurable construct, such as positive/negative affectivity, social desirability, or a theoretically irrelevant marker variable, is controlled), using confirmatory factor analysis (CFA) to control for a directly measured latent factor (again CMV is controlled via positive/negative affectivity, social desirability, or a theoretically irrelevant marker variable), controlling for an unmeasured latent factor, and the split-group technique.

## **FOUR APPROACHES TO CORRECTING FOR CMV**

*Partial Correlation Approaches.* In traditional partial correlation approaches, a measure that is an assumed source of method variance is set as a covariate in the statistical analysis of substantive variables (Podsakoff et al., 2003). The chosen covariate might be social desirability, positive affectivity, or negative affectivity, which are assumed to affect both predictor and criterion variables (Ganster, Hennessey, & Luthans, 1983; Williams & Anderson, 1994). A more recently proposed partial correlation technique has been detailed by Lindell and Whitney (2001). These authors suggest using the partial correlation approach, but with a covariate that is the best estimate of CMV present in a dataset. This best estimate is represented by the smallest correlation in the dataset between a substantive variable and a “marker variable” (Lindell & Whitney, 2001). Ideally, the marker is theoretically irrelevant to the substantive variables of interest and has a correlation with at least one of those variables that is close to zero. Lindell and

Whitney use the following equation to remove the variance shared with the marker from a substantive correlation:  $r_{Y_i.M} = r_{Y_i} - r_S / 1 - r_S$ , where  $r_{Y_i.M}$  is the partial correlation between Y and  $X_i$  controlling for CMV,  $r_{Y_i}$  is the observed correlation between Y and  $X_i$  suspected of being contaminated by CMV, and  $r_S$  is the smallest correlation between the marker variable and one of the substantive variables. The marker-adjusted correlation can also be corrected for disattenuation due to measurement error.

*Confirmatory Factor Analysis (CFA) Approach (with and without a measured method construct).* Williams et al. (2003) expanded Lindell and Whitney's work to a structural equations modeling context by building upon earlier work by Williams and colleagues (Williams & Anderson, 1994; Williams, Gavin, & Williams, 1996). Much like the partial correlation approaches, the CFA technique can be done using a directly measured latent variable (e.g., social desirability) or a theoretically irrelevant marker variable. The representative of CMV is modeled as a latent variable and, by altering various aspects of the measurement model specification, a series of models can be statistically compared to determine if method variance is present, the nature of the method variance (i.e., equal effects vs. unequal effects), its magnitude, and whether it is biasing observed correlations. When a measurable estimate of CMV is not available, Williams et al. (1989) have also suggested that one or more unmeasured latent method constructs can be specified as a means of partialling out variance shared among substantive indicators that is due neither to their substantive constructs nor to random error. Again, model comparison can be used to determine whether method variance is present and if it is biasing observed correlations.

*Split-Group Approach.* The split-group approach may only be used when individual-level data can be aggregated to a group level. This technique removes same-source bias, and thus CMV, because different respondents provide different data for the substantive variables (Ostroff et al., 2002; Podsakoff & Organ, 1986). That is, when data on independent and dependent variables is collected from groups, the groups can be split into subgroups in which one subgroup provides the data for the independent variable and the other provides the data for the dependent variable.

## METHOD

We simulated a five-variable model with four exogenous constructs (C1, C2, C3, and C4) and one endogenous construct (DV). Correlations among these were intended to represent substantive relationships. We modeled different true correlations for each exogenous-endogenous relationship, with C1-DV as strong ( $r = .70$ ), C2-DV as moderate ( $r = .40$ ), C3-DV as weak ( $r = .20$ ), and C4-DV as no ( $r = .00$ ) relationship. Each exogenous construct is measured by a four-item scale. The endogenous construct is measured by a five-item scale.

Based on this model, we simulated 18 datasets using a data generating program called DataSim (version 1.0.0, copyright 2002), created by Michael C. Sturman (2002). For three of the four techniques (the partial correlation, partial correlation marker method, and the split-group method), each dataset was created to be comprised of 300 respondents, which were organized into 50 groups of 6 individuals. For the CFA approaches, we generated data for a sample size of 420, organized into 70 groups of 6 individuals each, because this method requires a higher sample size for analysis. Correlations among the four exogenous variables were set at .25.

These datasets varied along the following conditions: (1) between group variance, (2) within-group variance, (3) common methods variance, and (4) random error variance. Between group variance represents the variance in the actual construct, or  $\rho$ , which is the true correlation between the independent and the dependent variable. In the datasets,  $\rho$  was determined based on the calculation of the other three dataset characteristics. The within-group variance is an indication of the amount of variance accounted for by the group; it can be conceptualized as group agreement. Common method variance is the variance due to the methods factor, and random variance is that which is accounted for by random error. In the 18 datasets, the total error variance was comprised of common method variance and random variance, and this was set at either 20% or 50%. Of that error variance, CMV was set at none (0% of total error variance), moderate (33% of total error variance), or high (66% of total error variance). When total error variance is removed, the remaining variance comprised between and within-group variance. Within-group variance was set at a zero (0%), moderate (33%), or high (66%) portion of the remaining variance. The between-group variance was then determined to be any variance remaining beyond the total error variance and the within-group variance.

## RESULTS

Using the observed exogenous-endogenous correlations in the 18 datasets, we examined the effectiveness of each correction technique for removing error. Error was defined as the absolute value of the difference between the “true” and observed (corrected) correlations. Error scores ranged from 0 to .93, with a mean of .18 and standard deviation of .18. Ultimately, we wanted to evaluate the relative utility of the various correction techniques; thus, we wanted to see how much error was reduced, on average, by each technique. However, we also recognized that the characteristics of the simulation would affect the magnitude and distribution of observed errors; thus, we conducted our analyses in two steps. The first step involved regressing the absolute errors on the following simulation characteristics: between group variance, within group variance, common method variance, random variance, true correlation, and whether the correlation was based on aggregate or individual-level data. Because characteristics of the simulation would likely interact with each other to affect the size of observed errors, we also considered up to three-way interactions between these six main effects. The resulting regression (results available upon request) explained 40% of the variance in the observed error scores.

The second step of our analyses entailed (a) saving the residuals from the above regression and (b) determining the extent to which the various correction techniques were associated with them. The residuals had a mean of zero, standard deviation of .14, a minimum of -.52, and a maximum of .60. We regressed the residual scores on a categorical variable representing the nine alternative correction techniques: no correction (labeled *none*), traditional partial correlation correction controlling for either a weak (labeled *partial(weak)*) or moderate (labeled *partial(moderate)*) estimate of CMV, the Lindell and Whitney approach (labeled  $r_{Yi.M}$ ), the Lindell and Whitney approach corrected for disattenuation (labeled  $r_{Yi.M}$  corrected), the CFA approach controlling for  $r_s$  (labeled *CFA.r*), the CFA approach controlling for an unmeasured latent factor (labeled *CFA.latent*), and the split group approach (labeled *split*). In the regression, we used the case of no correction (*none*) as the base case, so each correction was assessed relative to making no correction in data analysis. The results of the regression are shown in Table 1. The correction techniques, as a set, explained 32% of the variance of the residuals. Overall, the partial(weak), partial(moderate),  $r_{Yi.M}$ ,  $r_{Yi.M}$  corrected, and CFA.r techniques all yielded less

error than doing nothing. Alternatively, the CFA.latent and split techniques both actually produced more error on average than had no correction for common method variance been made.

We also wanted to examine our results for the presence of possible moderators. That is, we wanted to see if certain correction techniques were more effective given certain specific conditions of the simulation. For example, we explored if the effectiveness of the various correction techniques were moderated by the level of common method variance in the simulation. For this, we found no significant effects. Additionally, no effects were found for the level of random error in the simulation either. Aggregation had no effect, except for making the CFA.latent technique less accurate in the case of examining aggregated correlation coefficients. Significant moderating effects were found for the level of between-group variance, within-group variance, and the level of the “true” correlation in the simulation. Complete analyses are available from the authors; however, Table 1 reveals the mean effect on the residual of each correction technique. Note that, although we did find these moderating effects and they were statistically significant, there were no major sign changes in any of the results. This suggests that, overall, techniques that reduce error tend to reduce error across all of the conditions we explored, and that techniques that add error tend to add it across all of the conditions.

## DISCUSSION

While this is but a single study empirically investigating the utility of various CMV corrections, our results have important implications for research. Most notably, contrary to Podsakoff et al.’s (2003) warnings that this technique fails to control for some of the most powerful causes of CMV, the Lindell and Whitney approach appeared to be the most effective technique for eliminating CMV and producing estimates close to the “true” scores. Additionally, the split group technique and CFA controlling for an unmeasured latent method factor appear to create greater biases in the data than had no correction been made.

Though the Lindell and Whitney approach resulted in the most accurate corrections overall, there are some important practical implications that can influence its effectiveness when used with actual data. With this technique, the ultimate magnitude and accuracy of any given correction will be dependent upon the size of marker correlation ( $r_s$ ) and the extent to which the chosen value of  $r_s$  represents CMV as well as other sources of variance (e.g., true variance or random error). Although our results indicate that other characteristics of the dataset (e.g., amount of CMV or random error) do not substantially influence the accuracy of the corrections for this or the other techniques, it is possible that greater levels of random error, for instance, than were modeled here might influence the accuracy of this approach. That said, it is interesting to note that the values of  $r_s$  in the simulated data ranged from .01 to .93, meaning that in many datasets values of  $r_s$  far exceeded levels that Lindell and Whitney imply are appropriate. Fairly large values of  $r_s$  often produced highly accurate corrections as. However, as the size of  $r_s$  increased, the magnitude of the correlations corrected for disattenuation often reached meaningless levels.

The traditional partial correlation approach performed reasonably well in our simulated data, but the utility of this approach with actual data will also be dependent upon the extent to which a measurable representation of CMV is available to the researcher and the extent to which the researcher can be confident that the chosen covariate theoretically makes sense as an indicator of CMV. For instance, though personally sensitive survey items may be susceptible to bias from

social desirability, it is not reasonable to assume that all or even most items commonly included in organizational surveys are likely to be influenced by social desirability. It is also important to point out that, in actual data, the magnitudes of any partial correlation correction will always be dependent upon the size of both the exogenous-endogenous correlations and the exogenous/endogenous-control correlations.

Many of the implications and practical concerns discussed for the marker variable and traditional partial correlation approaches can also be applied to the CFA approach. However, there are two potential theoretical and practical advantages of the CFA approach over the partial correlation approaches. First, comparing the CFA approach to the Lindell and Whitney approach, CFA does not require the assumption of equivalent method effects across substantive indicators. Second, because the CFA approach includes a statistical test of bias via model comparison, researchers may be less likely to rely, and thus base their conclusions, on inappropriately corrected correlations (i.e., correlations not significantly biased by CMV).

Across the techniques examined here, our analyses suggest that researchers must use the most caution with the unmeasured latent method construct approach and the split group approach, both of which resulted in the least accurate corrections. One possible explanation for the poor results of the unmeasured latent method construct approach is that it requires highly parameterized models that may have problems with identification.

The poor results of the split-group technique, on the other hand, stem at least partially from two causes. First, we varied the amount of within-group variance across the datasets, and some datasets had very high and others had very low within group agreement. Without exception, the databases that resulted in completely nonsignificant corrected correlations for the split-group technique were those with low or no within group agreement. Second, whereas every other approach to correcting for CMV examined herein potentially removes multiple causes of CMV, the split-group method can only remove variance due to the same source. To the extent that other causes are present in the data, the split-group method will result in inaccurate corrections.

As a final observation, it is important to recognize that every approach had some influence on the observed correlations—even when no CMV was present in a given dataset. As such, it is inappropriate for researchers to use any of these approaches simply as a catch-all for dealing with CMV that *might* be present in their data. Rather, researchers should carefully consider and theoretically justify whether CMV is a plausible explanation for findings, and if it is, researchers should consider testing for the presence of bias (e.g., by using the CFA approach) in conjunction with whatever correction is ultimately used.

References are available from the authors upon request.

TABLE 1  
Average Effect of CMV Correction Techniques on Estimation Error

Technique	Over- all	"True" Correlation			Between-Group Variance			Within-Group Variance					
		.00	.20	.40	.70	17%	27%- 33%	50%- 53%	80%	0%	17%	27- 33%	53%
Partial (weak)	-.05**	-.04*	-.07**	-.06**	-.07*	-.06*	-.04*	-.04	-.03	-.03	-.06*	-.09*	
Partial (moderate)	-.07**	-.07*	-.08**	-.05*	-.10**	-.08**	-.07**	-.05*	-.05*	-.07*	-.09**	-.10**	
r <sub>yiM</sub>	-.07**	-.14**	-.03*	-.04*	-.09**	-.10**	-.05*	-.03	-.03*	-.05*	-.08**	-.14**	
r <sub>yiM</sub> corrected	-.08**	-.13**	-.04*	-.07**	-.12**	-.11**	-.06*	-.04	-.04*	-.06*	-.09**	-.17**	
CFA.r	-.04*	-.03	-.05*	-.04*	-.05*	-.02	-.03*	-.06*	-.05*	-.01	-.04*	-.04	
CFA.latent	.09**	.13**	.08**	.11	.20**	.11**	.03*	.03	.01	.06*	.14**	.17**	
Split	.16**	.11**	.14**	.26**	.10*	.16**	.17**	.22**	.21**	.11**	.12**	.21**	
Average  Error  (Standard Deviation)	.18 (.18)	.30 (.27)	.22 (.18)	.16 (.14)	.11 (.12)	.23 (.21)	.22 (.20)	.16 (.15)	.12 (.14)	.13 (.14)	.16 (.14)	.20 (.19)	.27 (.24)

Notes: Error was originally measured as |True Correlation - Estimated Correlation|. A regression was run to predict error, and the residuals saved. These residuals were predicted in a second regression, using the various CMV correction techniques. All cases above are compared against the case of using no correction. Negative values indicate that the observed scores are closer to the "true" scores.