Multivariate Behavioral Research

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hmbr20

On the Misconception of Multicollinearity in Detection of Moderating Effects: Multicollinearity Is Not Always Detrimental

Gwowen Shieh [a]

[a] National Chiao Tung University

Published online: 07 Jun 2010.

Multivariate Behavioral Research, 45:483–507, 2010

ISSN: 0027-3171 print/1532-7906 online
DOI: 10.1080/00273171.2010.483393

Psychology Press
Taylor & Francis Group

On the Misconception of Multicollinearity in Detection of Moderating Effects: Multicollinearity Is Not Always Detrimental

Gwowen Shieh
National Chiao Tung University

Due to its extensive applicability and computational ease, moderated multiple regression (MMR) has been widely employed to analyze interaction effects between 2 continuous predictor variables. Accordingly, considerable attention has been drawn toward the supposed multicollinearity problem between predictor variables and their cross-product term. This article attempts to clarify the misconception of multicollinearity in MMR studies. The counterintuitive yet beneficial effects of multicollinearity on the ability to detect moderator relationships are explored. Comprehensive treatments and numerical investigations are presented for the simplest interaction model and more complex three-predictor setting. The results provide critical insight that both helps avoid misleading interpretations and yields better understanding for the impact of intercorrelation among predictor variables in MMR analyses.

The use of moderated multiple regression (MMR) has become common across a wide variety of social science disciplines in the search for interaction effects. But despite its popularity, substantial concerns have been raised regarding the considerable difficulties of detecting moderation relationships that are strongly expected or theoretically supported. Numerous researchers have noted that the hypothesis tests of moderating effects often have low statistical power and yield erroneous

Correspondence concerning this article should be addressed to Gwowen Shieh, Department of Management Science, National Chiao Tung University, Hsinchu, Taiwan 30050, R.O.C. E-mail: gwshieh@mail.nctu.edu.tw

conclusions, impeding the theoretical development and scientific advancement of moderation research. In response to this problem, design considerations and model characteristics pertaining to power issues in MMR applications have been examined both conceptually and empirically. Notably, Aguinis (1995) identified prominent factors that attenuate statistical power and proposed practical solutions to low-power situations, especially for models with continuous moderators. On the other hand, Aguinis and Stone-Romero (1997) and Stone-Romero, Alliger, and Aguinis (1994) focused on the methodological artifacts and critical implications associated with statistical power of dichotomous moderators. Furthermore, the recent review by Aguinis, Beaty, Boik, and Pierce (2005) emphasized the importance of effect size and power in assessing moderating effects in the context of categorical moderators. In light of these discussions in the current literature, the responsible factors that stand out as being most crucial include sample size, magnitude of moderating effect, reliability of criterion and predictor variable scores, joint distribution of predictor variables, and intercorrelation of predictor variables.

In addition to the general treatment by Aguinis (1995) mentioned earlier, the multicollinearity problem in MMR has been examined by Cronbach (1987); Dunlap and Kemery (1987, 1988); Ganzach (1998); and Morris, Sherman, and Mansfield (1986), among others. It should be evident that the intercorrelation among the continuous predictor variables and their cross-product term is inevitably relevant to the detection of interaction in general. Hence, no single study of MMR with continuous variables will be adequate without considering the notion of multicollinearity. Accordingly, it is important to emphasize the distinction between essential and nonessential multicollinearity (Marquardt, 1980). Essential multicollinearity exists because of actual relationships between predictor variables, whereas the latter occurs merely due to the scaling or nonzero mean of predictor variable and can be removed by centering predictor variables. Related issues can be found in Kromrey and Foster-Johnson (1998), Smith and Sasaki (1979), and Tate (1984). It is generally known that other remedies exist for coping with multicollinearity as discussed in linear regression textbooks such as Cohen, Cohen, West and Aiken (2003) and Kutner, Nachtsheim, and Neter (2004). However, for clear understanding it is essential that researchers direct the subtle formulation and evaluation of moderating effects with sound theory and consider the delicate interrelationships and significance within the response and predictor variables. Specifically, a numerical example is provided in a later section to demonstrate the commonly used remedy of collecting additional data for alleviating the problem of multicollinearity. However, it does not yield the expected result in terms of increasing the ability to detect interaction effects.

In line with the foregoing concerns, Dunlap and Kemery (1988) examined the effects of both predictor reliabilities and predictor correlations on the statistical power of MMR. Their Monte Carlo simulation results showed that, as anticipated, the power to detect moderating effects is diminished by predictor unreliability. However, the corresponding empirical evidence gives rise to the surprising contention that the ability to detect interaction effects increases with increasing correlation between predictor variables. Because their discussions were focused more on the major issue of measurement error, and numerical findings were obtained from somehow limited settings in the context of two-predictor interaction models, Dunlap and Kemery (1988) did not provide insight into the counterintuitive power behavior in relation to multicollinearity diagnostics. It seems that this particular result has been overlooked in the literature and a further explanation that incorporates the notion of multicollinearity does not exist to our knowledge. Accordingly, it is of practical importance to assess whether this situation persists over a broader range of model configurations without the complication of unreliability.

In order to enhance the methodological integrity and fundamental usefulness of MMR, this article aims to explore the implications of intercorrelations among the continuous predictors and to account for misconception in the detection of moderating effects. In particular, the distinct power performance of the interactive models involving two predictor variables is presented to highlight the possible misapprehension when researchers apply heuristics learned from regular linear regression to MMR. Moreover, similar treatment and in-depth discussion are extended to the three-variable interaction model. For completeness, the Appendix summarizes the main results from the significance test of regression coefficients in the context of multiple linear regression with particular emphasis on the consideration of stochastic predictor variables. Informative figures and numerical results are presented to illustrate the essential features of MMR analyses.

TWO-PREDICTOR INTERACTION MODEL

Most MMR research has focused on the occurrence of interactive effects between two continuous predictor variables that are usually conceptualized in terms of the model

$$Y_i = \beta_I + \beta_X X_i + \beta_Z Z_i + \beta_{XZ} X_i Z_i + \varepsilon_i, \qquad (1)$$

where Y_i is the value of the response variable Y; X_i and Z_i are the known constants of the predictors X and Z; ε_i are *iid* $N(0, \sigma^2)$ random errors for $i = 1, \ldots, N$; and β_I, β_X, β_Z, and β_{XZ} are unknown parameters. The existence of the regression coefficient β_{XZ} associated with the cross-product term in Equation (1) indicates that the linear relationship between the criterion variable and predictor variable is dependent on the level of the other predictor variable. In contrast, the simple additive model without the multiplicative term

$$Y_i = \beta_I + \beta_X X_i + \beta_Z Z_i + \varepsilon_i$$

reveals that the association or strength between the response variable and each of the predictor variables is unaffected by or immaterial to the value of other predictor variables. The objective of MMR is to determine whether the underlying data structure can best be approximated by an additive or an interactive formulation. In practice, the detection of moderating effects is conducted with the partial F or partial t test for the hypothesis H_0: $\beta_{XZ} = 0$ versus H_1: $\beta_{XZ} \neq 0$ in the multiple linear regression framework.

It is generally known that the parameter estimation and hypothesis testing of multiple regression analysis can be plagued by the effects of multicollinearity. According to the fundamental properties of standard linear regression analysis presented in the Appendix, the estimated variances of the least squares coefficient estimators given in Equation (A5) are linked to the formal measure of variance inflation factor (VIF) for identifying the degree of multicollinearity. When a predictor variable has a strong linear association with other predictor variables, the associated VIF and variance estimate of regression coefficient estimator are excessively large. A commonly used rule of thumb is that a VIF of 10 or more is evidence of severe multicollinearity (Cohen et al., 2003, p. 423; Kutner et al., 2004, p. 387). Hence, the hypothesis testing of interaction effects is hampered and the power for detecting the moderation relationship is reduced because of the intercorrelation among the predictor variables.

Moreover, the adverse effects of multicollinearity on the linear regression analysis with the additive model are clearly apparent. Let $\hat{\beta}_X$ denote the least squares estimator of regression coefficient β_X, then the simple additive structure gives the following VIF of predictor variable X and estimated variance of $\hat{\beta}_X$:

$$VIF(X) = \frac{1}{1 - r^2} \text{ and } \hat{V}(\hat{\beta}_X) = \frac{\hat{\sigma}^2 \cdot VIF(X)}{S_X^2},$$

where $r = r(X, Z)$ is the Pearson product-moment correlation coefficient between the two predictor variables X and Z, $\hat{\sigma}^2$ is the usual unbiased estimator of σ^2, and $S_X^2 = \sum_{i=1}^{N}(X_i - \overline{X})^2$ is the corrected sum of squares with $\overline{X} = \sum_{i=1}^{N} X_i/N$. Similar results can be readily obtained for the second predictor variable Z. It is evident from the expressions just described that the degree of linear dependence between the two predictor variables measured by the simple correlation r has a significant influence on the multicollinearity index of VIF and the variance estimate $\hat{V}$ $(\hat{\beta}_X)$. The great simplicity of the additive model both makes it possible to convey the notion of multicollinearity without the burden of complex formulas and permits computational ease in empirical examination. For example, related implication and numerical illustration are well demonstrated in the acclaimed texts of Cohen et al. (2003, Sec. 10.5) and Kutner et al. (2004, Sec. 7.6). This reinforces the general perception and common practice that researchers should fully understand the intercorrelations among the predictor variables and carefully attend to the potential multicollinearity problem in a multiple regression analysis.

In view of the continuous characteristics of measurements X and Z, it is clear that the sample values and data characteristics in a study vary from one application to another. Accordingly, the value of simple correlation coefficient r represents only a realization of r over the whole range of $[-1, 1]$. Hence, it is of theoretical importance to investigate the overall impact of any underlying correlation between the two predictor variables on the various properties of MMR. In fact, the intercorrelation structure among the predictor variables is one of the inherent characteristics determined by the joint distribution of predictor variables, which in turn represents an indispensable artifact for detecting moderating effects. To extend the concept and applicability of MMR, it is more appropriate to employ the random regression or unconditional setup in which not only are values of the response variable for each participant available after the observations are made but the levels of predictor variables are also outcomes of the study. Thus the continuous predictor and moderator variables $\{(X_i, Z_i), i = 1, \ldots, N\}$ in Equation (1) are random variables with a joint probability distribution. This assumption is closely related to the consideration of stochastic regressors in econometrics. The impacts of the intercorrelation relationship on multicollinearity diagnostics and statistical features for identifying interaction effects are presented in the following analytical and numerical investigation.

Because of the complex nature of the random formulation under study, a complete theoretical solution is not feasible and the investigation is conducted in two stages. In the first stage, statistical derivations are carried out to gain an understanding of some specific phenomena for random regression models, subsuming the prescribed additive and interactive models and other MMR as special cases. The second stage is a large-scale simulation study in which pseudorandom data were generated with desired structural equations and then analyzed to determine the overall power behavior for discovering the main and interaction effects and unconditional performance of commonly used multicollinearity measures.

First, the corresponding important statistical features for identifying interaction effects and multicollinearity diagnostics with the extra complication of stochastic predictor variables are described in Equations (A7)–(A9) of the Appendix. The resulting formulas are difficult to comprehend in generic expressions; however, they allow various distributions for regressor variables to be treated as variations on a common theme and they serve to tie together the notions of moderation and correlation. Nevertheless, they contain essential information as to whether a given correlation structure reduces the power for detecting moderation relation whenever the distribution of predictor variables is available. Regarding the distributional assumptions of the associated predictor variables, it is common to assume that the two continuous predictor variables have a joint bivariate normal distribution in illustrative and theoretical treatments of MMR such as McClelland and Judd (1993), O'Connor (2006), and Shieh (2009). The bivariate normality assumption not only provides a useful situation in its own right but also has the advantage of naturally including the correlation between the two variables as a single free parameter. It is important to note that, although both X and Z are normally distributed, the interaction term XZ is obviously not a normal random variable. As mentioned earlier, joint distribution of the predictor variables is one of the deterministic factors of detecting moderating effects, and so it may distort statistical power analysis and lead to invalid conclusions if one mistakenly applies a multinormal setup to the regressors of MMR.

In the second stage of numerical examination, the prescribed interactive models with bivariate normal predictor variables are used as the base for Monte Carlo assessment. Without loss of generality, the two predictors (X, Z) are assumed to have a bivariate normal distribution with mean (0, 0), variance (1, 1), and correlation ρ ranging from -0.9 to 0.9 in increments of 0.1. Moreover, the power level is a function of regression coefficient β and error variance σ^2 through β/σ. Hence, the parameters are chosen as $\beta_I = \beta_X = \beta_Z = \beta_{XZ} = 0.25$ and $\sigma^2 = 1$. With sample size $N = 100$ and selected model configurations, the estimates of unconditional magnitudes are then computed through simulation of 10,000 replicate data sets. For each replicate, N sets of predictor variables are generated from the selected bivariate normal distribution. These values in turn determine the mean responses for generating N normal outcomes with the underlying linear regression model. Then the sample variance, test statistic, VIF, and regressor correlation matrix determinant (RCMD) are calculated. The simulated power is the proportion of the 10,000 replicates whose test statistic $|t|$ values exceed the critical value with significance level $\alpha = 0.05$. In addition, the overall estimates of variance, VIF, and RCMD are the arithmetic means of the corresponding 10,000 replicated values. All calculations were performed using programs written with SAS/IML (SAS Institute, 2008). Detailed numerical results of the simulation studies are reported in Table 1. Specifically, the simulated values of unconditional variance, power, and VIF associated with predictor X are denoted by $\nu(\hat{\beta}_X)$, $\pi(t_X)$ and $\phi(X)$, respectively, whereas the corresponding values for product term XZ are presented by $\nu(\hat{\beta}_{XZ})$, $\pi(t_{XZ})$ and $\phi(XZ)$. The overall RCMD is denoted by δ in Table 1 as well. Because predictors X and Z are interchangeable under bivariate normal distribution, the

TABLE 1
The Simulated Results of Two-Predictor Interaction Model
With $\beta_X = \beta_{XZ} = 0.25$ and $N = 100$

ρ	*Variance* $\nu(\hat{\beta}_X)$	*Power* $\pi(t_X)$	*VIF*[a] $\phi(X)$	*Variance* $\nu(\hat{\beta}_{XZ})$	*Power* $\pi(t_{XZ})$	*VIF*[a] $\phi(XZ)$	*RCMD*[b] δ
−0.9	0.0555	0.1858	5.4778	0.0067	0.8546	1.0641	0.1802
−0.8	0.0294	0.3078	2.8882	0.0074	0.8236	1.0641	0.3404
−0.7	0.0208	0.4094	2.0351	0.0082	0.7906	1.0633	0.4816
−0.6	0.0167	0.4898	1.6220	0.0089	0.7602	1.0634	0.6031
−0.5	0.0142	0.5532	1.3870	0.0095	0.7333	1.0630	0.7039
−0.4	0.0127	0.5978	1.2396	0.0102	0.7031	1.0642	0.7860
−0.3	0.0118	0.6317	1.1423	0.0108	0.6802	1.0625	0.8525
−0.2	0.0112	0.6541	1.0858	0.0112	0.6653	1.0628	0.8970
−0.1	0.0109	0.6651	1.0531	0.0116	0.6507	1.0643	0.9239
0	0.0107	0.6702	1.0423	0.0116	0.6500	1.0639	0.9337
0.1	0.0108	0.6661	1.0530	0.0115	0.6547	1.0630	0.9248
0.2	0.0112	0.6535	1.0845	0.0112	0.6632	1.0630	0.8978
0.3	0.0117	0.6324	1.1434	0.0107	0.6831	1.0629	0.8517
0.4	0.0127	0.5983	1.2361	0.0102	0.7038	1.0642	0.7877
0.5	0.0142	0.5524	1.3852	0.0096	0.7315	1.0623	0.7051
0.6	0.0166	0.4899	1.6232	0.0088	0.7609	1.0616	0.6030
0.7	0.0208	0.4094	2.0363	0.0082	0.7899	1.0640	0.4813
0.8	0.0294	0.3082	2.8870	0.0074	0.8253	1.0626	0.3408
0.9	0.0556	0.1855	5.4697	0.0068	0.8509	1.0643	0.1807

[a]Variance inflation factor. [b]Regressor correlation matrix determinant.

symmetric situations of predictor Z are omitted. For a concise visualization of the overall multicollinearity diagnostics with respect to the change of correlation ρ, Figure 1 depicts the relationship of simulated VIF for regressors X and XZ and RCMD with ρ. In addition, Figure 2 presents the plot of simulated power of t_X and t_{XZ} against ρ for the tests of main and interaction effects, respectively.

It is clear from Table 1 that the effect of positive and negative correlation ρ is symmetric on all seven measurements of variance, power, VIF, and determinant. In particular, Figure 1 reveals that the graphs of VIF measure $\phi(X)$ and determinant δ are symmetric with respect to $\rho = 0$ and the degrees of multicollinearity are increasing monotonous with increasing $|\rho|$. However, the VIF measure $\phi(XZ)$ remains almost constant. It should be noted that the unconditional variances have opposite patterns with respect to the correlation between X and Z. The overall $\nu(\hat{\beta}_X)$ is an increasing function of $|\rho|$, whereas $\nu(\hat{\beta}_{XZ})$ is decreasing with increasing magnitude of $|\rho|$. Moreover, the unconditional variance $\nu(\hat{\beta}_X)$ is larger than $\nu(\hat{\beta}_{XZ})$ for $|\rho| > 0.2$, and this situation is reversed for $|\rho| < 0.2$. The distinct behaviors of variances lead to power performance

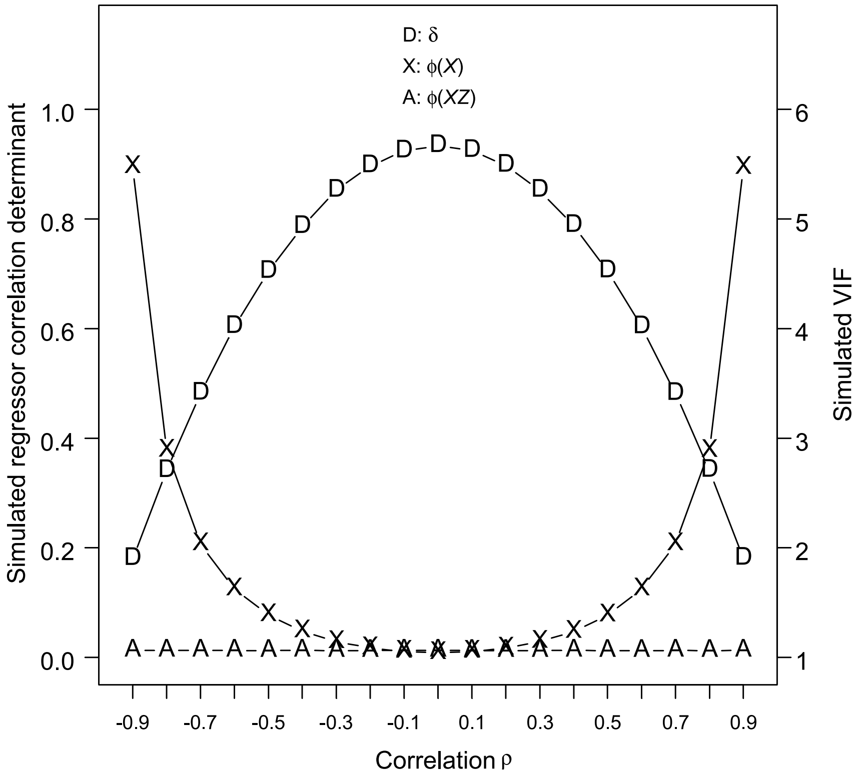

FIGURE 1 The simulated multicollinearity measures of two-predictor interaction model.

that is completely unexpected. As shown in Figure 2, the power function $\pi(t_X)$ decreases as the correlation becomes stronger, whereas the power of detecting interaction effects $\pi(t_{XZ})$ is essentially amplified for larger value of $|\rho|$. Hence, this particular exposition provides an obvious contradiction to the common impression that intercorrelation or multicollinearity between predictor variables is always detrimental to the power for detecting parameter effects. Consequently, researchers can make understandable but serious mistakes when they apply heuristics learned from simple additive models to MMR. Because the actual effect sizes of interaction terms in MMR applications are generally quite small, we also performed similar numerical computations for regression coefficients $\beta_I = \beta_X = \beta_Z = 0.25$, $\beta_{XZ} = 0.10$, and sample size $N = 250$, while all other factors remained constant. The corresponding results are presented in Table 2. Comparatively, the unconditional variances $\nu(\hat{\beta}_X)$ and $\nu(\hat{\beta}_{XZ})$ and power level $\pi(t_{XZ})$ are much smaller than those in Table 1. However, the

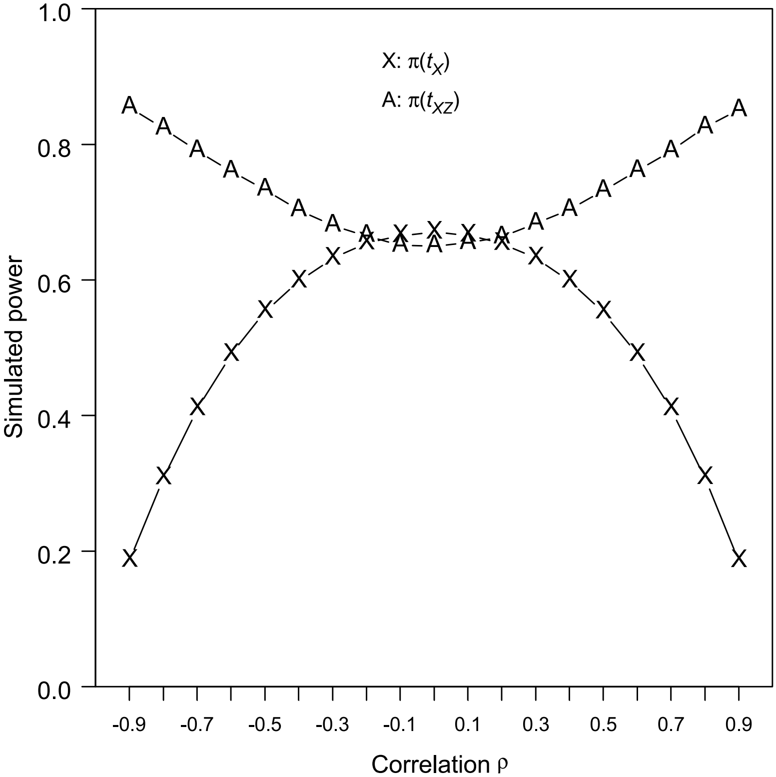

FIGURE 2 The simulated powers of two-predictor interaction model.

prescribed phenomena regarding their behavior relative to correlation ρ continue to exist in this case. In short, the advocated contention regarding the adverse relationship between multicollinearity and power in the literature for linear regression models does not generalize to MMR in a straightforward manner. The complex and yet important consequences of multiplicative components in MMR analyses will further be exemplified for three-predictor interaction models in the next section.

THREE-PREDICTOR INTERACTION MODEL

In view of the counterintuitive behavior in the most common procedure for detecting two-way interaction effects, it is prudent to extend the investigations to other widely useful MMR models. Particularly, the natural extension with

TABLE 2
The Simulated Results of Two-Predictor Interaction Model
With $\beta_X = 0.25$, $\beta_{XZ} = 0.10$, and $N = 250$

ρ	*Variance* $\nu(\hat{\beta}_X)$	*Power* $\pi(t_X)$	*VIF*[a] $\phi(X)$	*Variance* $\nu(\hat{\beta}_{XZ})$	*Power* $\pi(t_{XZ})$	*VIF*[a] $\phi(XZ)$	*RCMD*[b] δ
−0.9	0.0215	0.3992	5.3410	0.0024	0.5441	1.0247	0.1861
−0.8	0.0114	0.6481	2.8232	0.0026	0.5063	1.0246	0.3516
−0.7	0.0080	0.7941	1.9929	0.0029	0.4700	1.0247	0.4974
−0.6	0.0064	0.8747	1.5862	0.0032	0.4364	1.0240	0.6245
−0.5	0.0055	0.9197	1.3532	0.0034	0.4101	1.0245	0.7312
−0.4	0.0049	0.9443	1.2089	0.0037	0.3856	1.0241	0.8183
−0.3	0.0045	0.9583	1.1165	0.0039	0.3650	1.0248	0.8857
−0.2	0.0043	0.9660	1.0584	0.0041	0.3515	1.0249	0.9338
−0.1	0.0041	0.9706	1.0264	0.0042	0.3454	1.0242	0.9632
0	0.0041	0.9717	1.0162	0.0042	0.3416	1.0245	0.9727
0.1	0.0042	0.9704	1.0263	0.0042	0.3443	1.0244	0.9632
0.2	0.0043	0.9665	1.0584	0.0041	0.3529	1.0246	0.9341
0.3	0.0045	0.9584	1.1173	0.0039	0.3671	1.0244	0.8852
0.4	0.0049	0.9444	1.2093	0.0037	0.3861	1.0247	0.8178
0.5	0.0055	0.9194	1.3539	0.0034	0.4091	1.0248	0.7308
0.6	0.0064	0.8750	1.5865	0.0032	0.4376	1.0248	0.6240
0.7	0.0080	0.7943	1.9896	0.0029	0.4706	1.0242	0.4984
0.8	0.0113	0.6490	2.8185	0.0026	0.5057	1.0247	0.3522
0.9	0.0215	0.3988	5.3535	0.0024	0.5446	1.0244	0.1857

[a]Variance inflation factor. [b]Regressor correlation matrix determinant.

three predictor variables represents another important application of MMR in which the relation between the response variable Y and predictor variable X varies across levels of the other two predictor variables, Z and W, and their combinations. This results in the following three-predictor interaction model:

$$Y_i = \beta_I + \beta_X X_i + \beta_Z Z_i + \beta_W W_i + \beta_{XZ} X_i Z_i + \beta_{XW} X_i W_i + \beta_{ZW} Z_i W_i + \beta_{XZW} X_i Z_i W_i + \varepsilon_i, \tag{2}$$

where Y_i is the value of the response variable Y; X_i, Z_i, and W_i are the known constants of the predictors X, Z, and W; ε_i are *iid* $N(0, \sigma^2)$ random errors for $i = 1, \ldots, N$; and β_I, β_X, β_Z, β_W, β_{XZ}, β_{XW}, β_{ZW}, and β_{XZW} are unknown parameters. With the hierarchical or step-down approach, the regression coefficient β_{XZW} associated with the highest order product term of all three predictors XZW indicates the strength of the most essential moderating effect. On the other hand, the two-way interactions (β_{XZ}, β_{XW}, and β_{ZW}) and first-order effects (β_X, β_Z, and β_W) represent conditional effects that can be examined to facilitate the interpretation of the underlying complex interaction structure.

Readers can refer to Aiken and West (1991), Dawson and Richter (2006), and Jaccard and Turrisi (2003) for further details. To provide an insight into MMR research, the focus here is on the potential misunderstanding of the influence of multicollinearity within the context of three-predictor interaction model. Similar to the two-predictor case, a Monte Carlo simulation study was conducted to evaluate the influence of intercorrelations between predictor variables on the analysis of all first-, second- and third-order effects.

The empirical study involves multivariate normal predictor variables X, Z, and W with null means $\mu_X = \mu_Z = \mu_W = 0$, unit variance $\sigma_X^2 = \sigma_Z^2 = \sigma_W^2 = 1$, correlation $\text{Cor}(X, Z) = \rho$ ranging from -0.9 to 0.9 in increments of 0.1, and $\text{Cor}(X, W) = \text{Cor}(Z, W) = 0$. It should be clear from a theoretical standpoint that there are many situations with practical usefulness among sets of correlations. The designated correlation matrix of the three predictors represents merely a single possibility and serves the purpose well for demonstrating the concealed feature of MMR. Moreover, the model parameters in Equation (2) are chosen as $\beta_I = \beta_X = \beta_Z = \beta_W = \beta_{XZ} = \beta_{XW} = \beta_{ZW} = \beta_{XZW} = 0.25$, $\sigma^2 = 1$, and sample size $N = 100$. The simulation closely follows the previous numerical investigation in which the Monte Carlo integration procedure was implemented to determine the unconditional measurements through 10,000 replicate data sets.

The corresponding simulated results for main effects, two-way interactions, and three-way interaction are summarized in Tables 3–5, respectively. Due to the model's complexity, the resultant phenomenon can be made more comprehensible with the help of diagrams. The multicollinearity VIF measurements of regressors X, W, XZ, XW, and XZW, denoted by $\phi(X)$, $\phi(W)$, $\phi(XZ)$, $\phi(XW)$, and $\phi(XZW)$, respectively, and RCMD δ are depicted in Figure 3. Alternatively, the respective simulated power levels $\pi(t_X)$, $\pi(t_W)$, $\pi(t_{XZ})$, $\pi(t_{ZW})$, and $\pi(t_{XZW})$ of t tests t_X, t_W, t_{XZ}, t_{ZW}, and t_{XZW} are plotted in Figure 4. Because of the interchangeability between X and Z and XW and ZW, the results associated with regressors Z and ZW are not presented here. According to the visual information of Figure 3, all the diagrams of VIF values are concave whereas the RCMD curve is convex, but all are symmetric about $\rho = 0$. It follows from a simple guideline that multicollinearity is declared to exist whenever any VIF value is at least equal to 10. Thus, the resultant degrees of multicollinearity are not severe according to the reported magnitudes of VIF values. In contrast, the small δ values for $|\rho| > 0.5$ indicate that the degree of multicollinearity is considered problematic. The patterns of the VIF and RCMD diagnostics are unquestionably clear that the levels of intercorrelation among the regressors increase with the strength of correlation between the two predictors X and Z. Consequently, the heuristic about the adverse effects of multicollinearity would suggest that the corresponding estimated variance of regression coefficients should be inflated and power of the resulting test

TABLE 3
The Simulated Results for X and W of Three-Predictor Interaction Model With $\beta_X = \beta_W = 0.25$ and $N = 100$

ρ	*Variance* $\nu(\hat{\beta}_X)$	*Power* $\pi(t_X)$	*VIF*[a] $\phi(X)$	*Variance* $\nu(\hat{\beta}_W)$	*Power* $\pi(t_W)$	*VIF*[a] $\phi(W)$
−0.9	0.0593	0.1772	5.8344	0.0172	0.4809	1.6646
−0.8	0.0315	0.2914	3.0941	0.0164	0.4986	1.5907
−0.7	0.0224	0.3866	2.1862	0.0156	0.5188	1.5103
−0.6	0.0178	0.4641	1.7464	0.0148	0.5397	1.4297
−0.5	0.0154	0.5218	1.4945	0.0139	0.5638	1.3489
−0.4	0.0138	0.5646	1.3411	0.0132	0.5853	1.2794
−0.3	0.0127	0.5989	1.2387	0.0125	0.6062	1.2171
−0.2	0.0122	0.6184	1.1777	0.0121	0.6216	1.1718
−0.1	0.0118	0.6323	1.1430	0.0118	0.6319	1.1427
0	0.0117	0.6358	1.1333	0.0117	0.6354	1.1323
0.1	0.0118	0.6304	1.1439	0.0118	0.6310	1.1420
0.2	0.0121	0.6189	1.1760	0.0120	0.6231	1.1696
0.3	0.0128	0.5970	1.2396	0.0126	0.6046	1.2181
0.4	0.0138	0.5664	1.3409	0.0132	0.5858	1.2782
0.5	0.0153	0.5237	1.4920	0.0139	0.5631	1.3518
0.6	0.0179	0.4635	1.7449	0.0147	0.5401	1.4323
0.7	0.0223	0.3871	2.1868	0.0156	0.5192	1.5109
0.8	0.0315	0.2912	3.0974	0.0164	0.4997	1.5915
0.9	0.0594	0.1770	5.8503	0.0172	0.4802	1.6667

[a] Variance inflation factor.

of main effects, two-way interactions, or three-way interaction will decline as the only present pairwise correlation ρ of X and Z increases in absolute size. The results show that the general notion is applicable only to the cases associated with regressors X, W, and cross-product XW. In other words, the unconditional estimated variances $\nu(\hat{\beta}_X)$, $\nu(\hat{\beta}_W)$, and $\nu(\hat{\beta}_{XW})$ are convex functions of correlation ρ, and conversely, power levels $\pi(t_X)$, $\pi(t_W)$, and $\pi(t_{ZW})$ are concave with respect to correlation ρ. Nonetheless, the conventional account does not apply to the other two regressors in terms of product terms XZ and XZW. Surprisingly, the two variance estimates $\nu(\hat{\beta}_{XZ})$ and $\nu(\hat{\beta}_{XZW})$ are concave with respect to ρ, and in turn, the respective power functions $\pi(t_{XZ})$ and $\pi(t_{XZW})$ are convex, as shown in Figure 4. Thus, the established guidance about the detrimental impact of multicollinearity in the context additive multiple regression is not completely applicable to interaction models. As in the previous case of a two-predictor interaction model, the empirical investigation was extended to the setting with $\beta_I = \beta_X = \beta_Z = \beta_W = 0.25$, $\beta_{XZ} = \beta_{XW} = \beta_{ZW} = 0.15$, $\beta_{XZW} = 0.10$, and sample size $N = 250$. According the results summarized in Tables 6–8, it is clear that the general contention described

TABLE 4
The Simulated Results for XZ and XW of Three-Predictor Interaction Model With $\beta_{XZ} = \beta_{XW} = 0.25$ and $N = 100$

ρ	*Variance* $\nu(\hat{\beta}_{XZ})$	*Power* $\pi(t_{XZ})$	*VIF*[a] $\phi(XZ)$	*Variance* $\nu(\hat{\beta}_{XW})$	*Power* $\pi(t_{XW})$	*VIF*[a] $\phi(XW)$
−0.9	0.0080	0.7994	1.2692	0.0663	0.1696	6.3426
−0.8	0.0087	0.7666	1.2701	0.0354	0.2762	3.3766
−0.7	0.0096	0.7323	1.2693	0.0251	0.3654	2.3758
−0.6	0.0103	0.7026	1.2588	0.0202	0.4353	1.8955
−0.5	0.0112	0.6685	1.2550	0.0174	0.4882	1.6215
−0.4	0.0120	0.6419	1.2493	0.0157	0.5285	1.4525
−0.3	0.0125	0.6232	1.2407	0.0145	0.5596	1.3417
−0.2	0.0129	0.6076	1.2350	0.0138	0.5785	1.2786
−0.1	0.0132	0.5986	1.2291	0.0134	0.5905	1.2428
0	0.0133	0.5936	1.2282	0.0133	0.5937	1.2286
0.1	0.0132	0.5964	1.2329	0.0135	0.5893	1.2429
0.2	0.0130	0.6053	1.2333	0.0138	0.5792	1.2743
0.3	0.0126	0.6209	1.2398	0.0146	0.5562	1.3434
0.4	0.0118	0.6452	1.2497	0.0156	0.5304	1.4505
0.5	0.0112	0.6704	1.2562	0.0174	0.4901	1.6211
0.6	0.0105	0.6974	1.2591	0.0201	0.4370	1.8964
0.7	0.0096	0.7313	1.2694	0.0251	0.3656	2.3778
0.8	0.0087	0.7667	1.2747	0.0352	0.2774	3.3689
0.9	0.0080	0.7989	1.2692	0.0666	0.1692	6.3851

[a]Variance inflation factor.

earlier can still apply in this situation with smaller effect size. Although these empirical examinations depend exclusively on simulation results, the assessments of the three-predictor interaction formulation illustrate the advocated caution and unfavorable perception of intercorrelations among predictor variables should not be applied indiscriminately. More important, the positive influence of correlation ρ on the detection of a three-way moderating effect raises a practical concern for MMR researchers to reevaluate the underlying predictor interrelationships and their impact on model selection and inference.

NUMERICAL EXAMPLE

In addition to the detailed empirical investigations employing Monte Carlo simulation study, it is instructive to exemplify the impact of multicollinearity on the detection of three-way interactions that might be encountered in applied work. The study of the importance of relationship in Kwong and Leung (2002) is used as an illustrative context. In that study they examined the compensatory effect

TABLE 5
The Simulated Results for *XZW* of Three-Predictor Interaction Model With $\beta_{XZW} = 0.25$ and $N = 100$

ρ	*Variance* $\nu(\hat{\beta}_{XZW})$	*Power* $\pi(t_{XZW})$	*VIF*[a] $\phi(XZW)$	*RCMD*[b] δ
−0.9	0.0101	0.7237	2.0587	0.0152
−0.8	0.0110	0.6936	1.9699	0.0553
−0.7	0.0121	0.6593	1.8747	0.1154
−0.6	0.0130	0.6302	1.7652	0.1896
−0.5	0.0140	0.6009	1.6598	0.2735
−0.4	0.0149	0.5743	1.5712	0.3555
−0.3	0.0156	0.5588	1.4848	0.4355
−0.2	0.0160	0.5457	1.4225	0.4982
−0.1	0.0163	0.5365	1.3848	0.5405
0	0.0165	0.5305	1.3702	0.5552
0.1	0.0164	0.5355	1.3894	0.5382
0.2	0.0160	0.5429	1.4195	0.5011
0.3	0.0156	0.5560	1.4829	0.4359
0.4	0.0148	0.5770	1.5698	0.3566
0.5	0.0140	0.6009	1.6671	0.2722
0.6	0.0131	0.6273	1.7653	0.1900
0.7	0.0121	0.6594	1.8730	0.1156
0.8	0.0110	0.6940	1.9731	0.0548
0.9	0.0101	0.7244	2.0586	0.0150

[a]Variance inflation factor. [b]Regressor correlation matrix determinant.

between procedural justice and outcome favorability in determining people's reaction to a decision. Given the compensatory effect, procedural fairness has a particularly strong and positive impact on people's response to low outcomes. However, they argued that the compensatory effect is conditional upon other contextual variables and studied the three-way interaction in which the perceived importance of the relationship between people moderates the compensatory effect of procedural justice. They tested the hypothesis that the tendency for procedural justice to have a stronger and more positive impact on people's response when outcome is low versus high should be more pronounced for an important relationship than for an unimportant relationship. The study concluded that the interaction effect is operative only when the relationship with the other party is important to that person.

For the purpose of demonstration, the summary statistics and analysis results presented in Tables 1 and 2 of Kwong and Leung (2002) were utilized to generate the two hypothetical data sets reported in Table 9. According to the formulation of Kwong and Leung, the criterion variable (Y) represents the measurement of

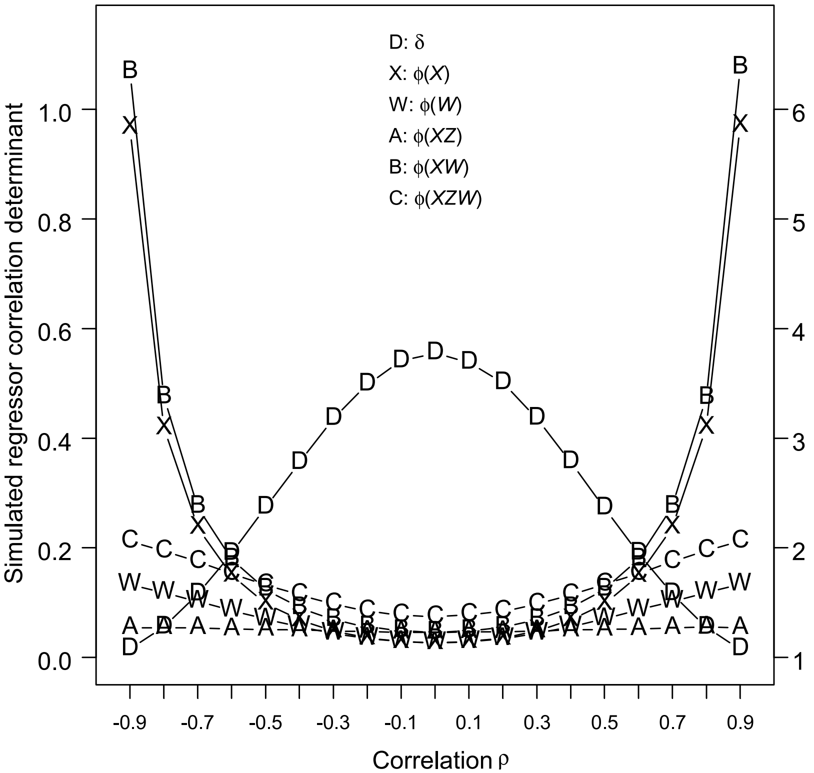

FIGURE 3 The simulated multicollinearity measures of three-predictor interaction model.

feeling or happiness, and the three predictor variables are interactional justice (X), outcome favorability (Z), and prior closeness (W). As noted in Aiken and West (1991, p. 36), the so-called nonessential multicollinearity can be removed by centering variables. Hence, the observed values of the three predictors in Table 9 were mean-centered in the following MMR analyses. With the 30 observations in Data 1, the simple correlations are $r(X, Z) = 0.4883$, $r(X, W) = 0.3541$, and $r(Z, W) = 0.2605$. The sample data was analyzed with a three-way interaction regression model. We are particularly concerned with the interaction term XZW, and the resulting test statistic is $t_{XZW} = -2.1873$ with p value $= .0396$. Hence, the test of three-way interaction H_0: $\beta_{XZW} = 0$ can be rejected at the significance level $\alpha = 0.05$. However, close examination of the variance inflation factor associated with the cross-product term XZW shows that $VIF(XZW) = 11.94$ and regressor correlation matrix determinant RCMD $=$ 0.0098. In practice, the VIF values in excess of 10 or the quantities RCMD

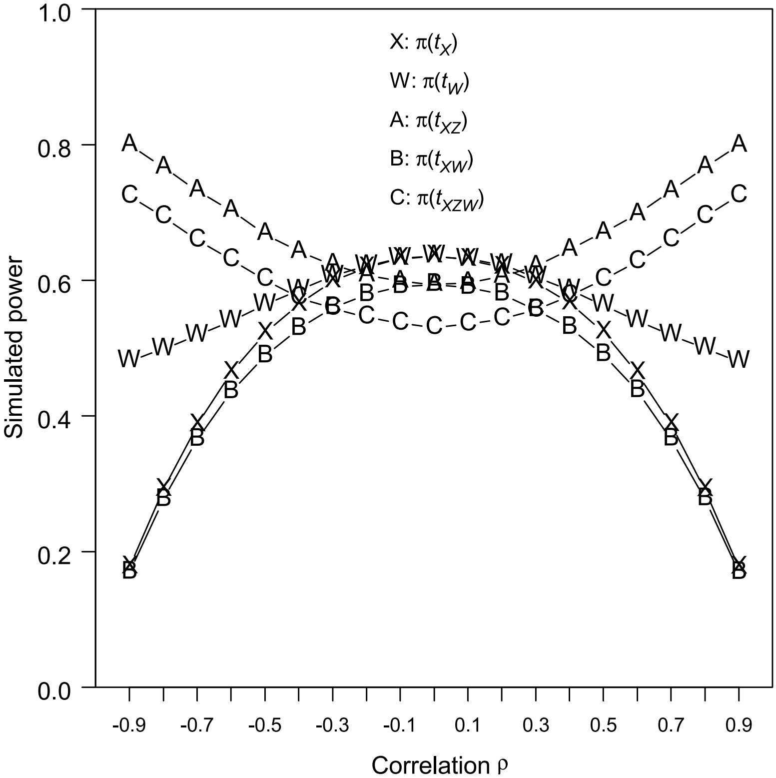

FIGURE 4 The simulated powers of three-predictor interaction model.

close to 0 are considered problematic. In these circumstances, the common procedure is to consider approaches to solving the problem of multicollinearity before concluding that there is sufficient evidence to indicate an interaction. Accordingly, the collection of additional data provides a feasible solution and is commonly recommended. With the additional 20 observations presented in Data 2 of Table 9, the detection of three-way interaction was reanalyzed with a total of sample size $N = 50$. In this case, the three pairwise correlations are $r(X, Z) = 0.4799$, $r(X, W) = 0.2308$, and $r(Z, W) = 0.1868$. The magnitudes of these correlations are less than those calculated with Data 1. Moreover, the multicollinearity index VIF reduced to $VIF(XZW) = 2.99$, whereas regressor correlation matrix determinant changed into RCMD $= 0.1083$. Thus, the severity of multicollinearity is alleviated to some extent as intended by the inclusion of extra samples. However, the resulting test statistic for the interaction effect is $t_{XZW} = -1.9104$ and the corresponding p value $= .0629$.

TABLE 6
The Simulated Results for X and W of Three-Predictor Interaction Model
With $\beta_X = \beta_W = 0.25$ and $N = 250$

ρ	*Variance* $\nu(\hat{\beta}_X)$	*Power* $\pi(t_X)$	*VIF*[a] $\phi(X)$	*Variance* $\nu(\hat{\beta}_W)$	*Power* $\pi(t_W)$	*VIF*[a] $\phi(W)$
−0.9	0.0221	0.3909	5.4768	0.0062	0.8831	1.5380
−0.8	0.0117	0.6377	2.8933	0.0060	0.8959	1.4725
−0.7	0.0083	0.7830	2.0471	0.0057	0.9093	1.4026
−0.6	0.0066	0.8659	1.6328	0.0054	0.9228	1.3302
−0.5	0.0056	0.9117	1.3923	0.0051	0.9354	1.2577
−0.4	0.0050	0.9382	1.2469	0.0048	0.9459	1.1923
−0.3	0.0047	0.9531	1.1515	0.0046	0.9556	1.1335
−0.2	0.0044	0.9618	1.0921	0.0044	0.9620	1.0884
−0.1	0.0043	0.9659	1.0596	0.0043	0.9659	1.0599
0	0.0042	0.9673	1.0497	0.0042	0.9674	1.0493
0.1	0.0043	0.9661	1.0600	0.0043	0.9660	1.0600
0.2	0.0044	0.9616	1.0921	0.0044	0.9621	1.0880
0.3	0.0047	0.9531	1.1523	0.0046	0.9552	1.1337
0.4	0.0050	0.9380	1.2456	0.0048	0.9465	1.1923
0.5	0.0056	0.9119	1.3936	0.0051	0.9350	1.2579
0.6	0.0066	0.8653	1.6323	0.0054	0.9223	1.3315
0.7	0.0083	0.7830	2.0484	0.0057	0.9094	1.4023
0.8	0.0117	0.6375	2.8914	0.0060	0.8958	1.4723
0.9	0.0221	0.3910	5.4838	0.0062	0.8829	1.5405

[a]Variance inflation factor.

Unfortunately, we are unable to claim there is a significant moderation effect at a 0.05 level of significance for the expanded data. This numerical illustration contradicts the preconceived notion about the attempt to resolve the threat of multicollinearity, thereby increasing the precision of estimates and improving the ability to detect interaction effects. Such finding provides a better understanding and demonstration on the diverse impact of predictor intercorrelations in MMR applications.

IMPLICATION FOR MODERATION ANALYSIS

Unlike the typical results of regression analyses, the prescribed simulation study and numerical example reveal a contrasting and positive impact of predictor intercorrelations on the detection of moderating effects. Researchers using MMR should be aware of this tendency of power for the detection of moderator effects to be lost at the expense of overemphasis on mitigation of multicollinearity between predictor variables. From the methodological standpoint, the techniques of

TABLE 7
The Simulated Results for *XZ* and *XW* of Three-Predictor Interaction Model With $\beta_{XZ} = \beta_{XW} = 0.15$ and $N = 250$

ρ	*Variance* $\nu(\hat{\beta}_{XZ})$	*Power* $\pi(t_{XZ})$	*VIF*[a] $\phi(XZ)$	*Variance* $\nu(\hat{\beta}_{XW})$	*Power* $\pi(t_{XW})$	*VIF*[a] $\phi(XW)$
−0.9	0.0026	0.8371	1.1045	0.0232	0.1694	5.6767
−0.8	0.0028	0.8032	1.1053	0.0123	0.2780	3.0023
−0.7	0.0031	0.7667	1.1043	0.0087	0.3693	2.1172
−0.6	0.0034	0.7336	1.1029	0.0069	0.4437	1.6888
−0.5	0.0037	0.6982	1.0994	0.0059	0.5007	1.4407
−0.4	0.0039	0.6718	1.0966	0.0053	0.5430	1.2904
−0.3	0.0042	0.6463	1.0927	0.0049	0.5772	1.2920
−0.2	0.0043	0.6282	1.0889	0.0047	0.5976	1.1313
−0.1	0.0044	0.6187	1.0868	0.0046	0.6089	1.0985
0	0.0045	0.6136	1.0888	0.0045	0.6141	1.0876
0.1	0.0045	0.6167	1.0889	0.0045	0.6099	1.0988
0.2	0.0044	0.6278	1.0907	0.0047	0.5968	1.1306
0.3	0.0041	0.6489	1.0915	0.0049	0.5750	1.1936
0.4	0.0039	0.6709	1.0946	0.0053	0.5440	1.2898
0.5	0.0037	0.7001	1.0990	0.0060	0.4995	1.4416
0.6	0.0034	0.7315	1.1000	0.0070	0.4420	1.6901
0.7	0.0031	0.7680	1.1040	0.0087	0.3697	2.1185
0.8	0.0029	0.8012	1.1064	0.0123	0.2779	2.9962
0.9	0.0026	0.8365	1.1053	0.0231	0.1696	5.6860

[a] Variance inflation factor.

multiple regression and other multivariate methods were developed to synthesize the complex information of correlated data in the first place. It seems paradoxical that the common practice has been overwhelmingly prone to remove or diminish intercorrelation and multicollinearity among variables, whereas the advanced methodologies are supposed to fully account for their intertwined structure in order to help advance social science theory.

CONCLUSION

It is well known that multicollinearity is closely related to the popular statistical tool of multiple linear regression. Hence, practitioners in applied research must become conversant with various diagnostic procedures for identifying, reducing, or removing the cause and threat of multicollinearity. The simplest MMR is essentially a special case of multiple linear regression that allows particularly the relation between the response variable and a predictor variable to depend on the level of another predictor variable. The basic rationale of moderation can

TABLE 8
The Simulated Results for *XZW* of Three-Predictor Interaction Model With $\beta_{XZW} = 0.10$ and $N = 250$

ρ	*Variance* $\nu(\hat{\beta}_{XZW})$	*Power* $\pi(t_{XZW})$	*VIF*[a] $\phi(XZW)$	*RCMD*[b] δ
−0.9	0.0029	0.4856	1.6918	0.0199
−0.8	0.0032	0.4535	1.6187	0.0740
−0.7	0.0035	0.4200	1.5440	0.1546
−0.6	0.0038	0.3927	1.4638	0.2551
−0.5	0.0041	0.3664	1.3817	0.3700
−0.4	0.0044	0.3476	1.3065	0.4862
−0.3	0.0046	0.3321	1.2394	0.5991
−0.2	0.0048	0.3196	1.1854	0.6935
−0.1	0.0049	0.3127	1.1522	0.7554
0	0.0050	0.3110	1.1425	0.7756
0.1	0.0050	0.3125	1.1545	0.7535
0.2	0.0048	0.3195	1.1864	0.6931
0.3	0.0046	0.3336	1.2382	0.5991
0.4	0.0044	0.3472	1.3040	0.4883
0.5	0.0041	0.3669	1.3806	0.3701
0.6	0.0038	0.3907	1.4602	0.2557
0.7	0.0035	0.4222	1.5425	0.1545
0.8	0.0032	0.4504	1.6206	0.0740
0.9	0.0029	0.4866	1.6914	0.0198

[a]Variance inflation factor. [b]Regressor correlation matrix determinant.

be readily extended to three-way interactions and more complex situations. In view of the apparent intercorrelated structure between the predictor variables and their combined higher order or cross-product terms in interaction models, the supposed adverse effects associated with high or extreme multicollinearity are often encountered in many MMR applications. Unfortunately, it is subject to serious misunderstanding that predictor intercorrelations incur nothing but harm to the detection of moderation or interaction effects in MMR study.

This article focuses on the two most fundamental MMR models with two- and three-predictor interaction effects and explores the impact of intercorrelations on the multicollinearity diagnostics and power in testing for main and interaction effects under the convenient distributional assumption of bivariate or multivariate normal predictors. The extensive empirical results of Monte Carlo simulation studies showed that the power of detecting interaction effects may increase with greater correlation between predictor variables when all other factors are fixed. Hence, the detrimental effects of multicollinearity associated with additive multiple linear regression are not necessarily present with MMR analysis. Regarding

TABLE 9
Hypothetical Data Sets

Y	*X*	*Z*	*W*	*Y*	*X*	*Z*	*W*	*Y*	*X*	*Z*	*W*
Data 1 ($N = 30$)											
2.02	3.2	2.1	2.6	3.19	2.6	2.1	1.6	3.93	3.7	2.8	4.7
3.32	3.5	3.2	3.1	3.76	2.8	1.9	1.7	1.96	4.0	2.9	3.6
2.63	3.6	3.3	1.3	4.64	3.8	3.3	2.4	3.78	3.3	2.4	0.3
1.43	4.3	4.1	5.3	2.21	3.6	1.6	2.3	1.69	2.1	2.9	1.7
3.40	2.9	2.5	3.5	5.07	3.6	4.1	3.0	5.47	3.9	2.7	3.8
2.24	4.0	2.5	3.8	1.87	3.8	3.1	3.2	3.03	2.6	3.4	2.8
1.65	1.4	0.6	2.9	4.24	4.3	2.8	2.7	2.54	2.9	2.1	3.2
3.39	4.6	5.1	4.9	3.88	2.8	2.9	3.9	3.69	3.6	3.5	1.7
1.98	2.2	4.0	2.6	3.20	2.7	2.7	3.3	4.72	4.2	3.4	2.3
3.02	3.3	2.5	5.9	3.21	0.8	2.6	1.8	1.01	2.4	1.3	2.6
Data 2 ($N = 20$)											
3.62	4.2	3.6	2.3	3.99	2.5	1.4	1.1	3.76	2.1	1.7	3.7
1.35	2.9	1.2	0.9	5.34	4.6	4.1	2.1	3.68	2.9	3.2	4.4
4.08	3.8	3.1	1.7	3.52	2.2	2.9	1.9	0.99	3.0	1.7	3.4
1.57	2.2	2.5	1.5	2.24	2.1	2.4	2.9	2.86	2.5	2.9	3.9
3.39	3.7	2.6	3.6	2.95	2.7	2.0	3.5	3.30	2.4	3.1	2.6
5.12	5.2	3.9	2.3	3.42	2.7	4.0	3.4	2.13	2.6	2.5	3.0
3.31	2.5	3.7	1.9	2.66	4.0	2.4	4.7				

the distributional configuration of predictor variables, normality is of course not the only situation of practical interest. There are also many useful assumptions to consider for the continuous predictor variables. More important, additional Monte Carlo simulations confirmed that the emphasized counterintuitive phenomenon is not unique to the normality assumption of the predictor variables. In view of the indispensable role of the joint distribution of predictors, researchers should make a comprehensive appraisal of the underlying data characteristics and their impact on statistical power for the detection of moderating effects. Given the complex interrelationships that exist among predictor variables and cross-product terms in MMR studies, it is important to reorient the general idea about the presence of multicollinearity because the detection of interaction effects need not be hindered by increased correlation between predictor variables.

ACKNOWLEDGMENTS

I thank the editor, Dr. Joseph Lee Rodgers, and the two anonymous reviewers for their valuable comments on earlier drafts of the article.

REFERENCES

Aguinis, H. (1995). Statistical power problems with moderated multiple regression in management research. *Journal of Management, 21,* 1141–1158.

Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology, 90,* 94–107.

Aguinis, H., & Stone-Romero, E. F. (1997). Methodological artifacts in moderated multiple regression and their effects on statistical power. *Journal of Applied Psychology, 82,* 192–206.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Cronbach, L. J. (1987). Statistical tests for moderator variables: Flaws in analyses recently proposed. *Psychological Bulletin, 102,* 414–417.

Dawson, J. F., & Richter, A. W. (2006). Probing three-way interactions in moderated multiple regression: Development and application of a slope difference test. *Journal of Applied Psychology, 91,* 917–926.

Dunlap, W. P., & Kemery, E. R. (1987). Failure to detect moderating effects: Is multicollinearity the problem? *Psychological Bulletin, 102,* 418–420.

Dunlap, W. P., & Kemery, E. R. (1988). Effects of predictor intercorrelations and reliabilities on moderated multiple regression. *Organizational Behavior and Human Decision Processes, 41,* 248–258.

Ganzach, Y. (1998). Nonlinearity, multicollinearity and the probability of type II error in detecting interaction. *Journal of Management, 24,* 615–622.

Jaccard, J., & Turrisi, R. (2003). *Interaction effects in multiple regression* (2nd ed.). Thousand Oaks, CA: Sage.

Kromrey, J. D., & Foster-Johnson, L. (1998). Mean centering in moderated multiple regression: Much ado about nothing. *Educational and Psychological Measurement, 58,* 42–67.

Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). *Applied linear regression models* (4th ed.). New York: McGraw-Hill.

Kwong, J. Y. Y., & Leung, K. (2002). A moderator of the interaction effect of procedural justice and outcome favorability: Importance of the relationship. *Organizational Behavior and Human Decision Processes, 87,* 278–299.

Marquardt, D. W. (1980). You should standardize the predictor variables in your regression models. *Journal of the American Statistical Association, 75,* 87–91.

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin, 114,* 376–390.

Morris, J. H., Sherman, J. D., & Mansfield, E. R. (1986). Failures to detect moderating effects with ordinary least squares-moderated multiple regression: Some reasons and a remedy. *Psychological Bulletin, 99,* 282–288.

O'Connor, B. P. (2006). Programs for problems created by continuous variable distributions in moderated multiple regression. *Organizational Research Methods, 9,* 554–567.

SAS Institute. (2008). *SAS/IML user's guide*, *Version 9.2*. Cary, NC: Author.

Shieh, G. (2009). Detecting interaction effects in moderated multiple regression with continuous variables: Power and sample size considerations. *Organizational Research Methods, 12,* 510–528.

Smith, K. W., & Sasaki, M. S. (1979). Decreasing multicollinearity: A method for models with multiplicative functions. *Sociological Methods & Research, 8,* 35–56.

Stone-Romero, E. F., Alliger, G. M., & Aguinis, H. (1994). Type II error problems in the use of moderated multiple regression for the detection of moderating effects of dichotomous variables. *Journal of Management, 20,* 167–178.

Tate, R. (1984). Limitations of centering for interactive models. *Sociological Methods & Research, 13,* 251–271.

APPENDIX

Fundamental Results of Random Regression Models

Consider the standard multiple linear regression model with dependent variable Y and all the levels of p independent variables $X_1, \ldots, X_p$ fixed a priori:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{A1}$$

where $\mathbf{Y} = (Y_1, \ldots, Y_N)^{\mathrm{T}}$, Y_i is the value of the dependent variable Y; $\mathbf{X} = (\mathbf{1}_N, \mathbf{X}_D)$ with $\mathbf{1}_N$ is the $N \times 1$ vector of all 1s, $\mathbf{X}_D = (\mathbf{X}_1, \ldots, \mathbf{X}_N)^{\mathrm{T}}$ is often called the design matrix, $\mathbf{X}_i = (X_{i1}, \ldots, X_{ip})^{\mathrm{T}}$, $X_{i1}, \ldots, X_{ip}$ are the known constants of the p independent variables for $i = 1, \ldots, N$; $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^{\mathrm{T}}$ with $\beta_0, \beta_1, \ldots, \beta_p$ are unknown parameters; and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N)^{\mathrm{T}}$ with ε_i are *iid* $N(0, \sigma^2)$ random variables for $i = 1, \ldots, N$.

Frequently, the inferences are concerned mainly with the regression coefficients $\boldsymbol{\beta}_1 = (\beta_1, \ldots, \beta_p)^{\mathrm{T}}$ in Equation (A1) and the corresponding ordinary least squares estimator is $\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}_C^{\mathrm{T}}\mathbf{X}_C)^{-1}\mathbf{X}_C^{\mathrm{T}}\mathbf{Y}$, where $\mathbf{X}_C = (\mathbf{I}_N - \mathbf{J}/N)\mathbf{X}_D$ is the centered form of $\mathbf{X}_D$, $\mathbf{I}_N$ is the identity matrix of dimension N, and $\mathbf{J}$ is the $N \times N$ square matrix of 1s. With this formulation, it is easily seen that

$$\hat{\boldsymbol{\beta}}_1 | \mathbf{X}_D \sim N_p(\boldsymbol{\beta}_1, \sigma^2 \mathbf{S}_X^{-1}),$$

where $\mathbf{S}_X = \mathbf{X}_C^{\mathrm{T}}\mathbf{X}_C$. Note that $\hat{\sigma}^2 = SSE/(N - p - 1)$ is the usual unbiased estimator of σ^2 and SSE/σ^2 is distributed as $\chi^2(N - p - 1)$, a chi-square distribution with $N - p - 1$ degrees of freedom and is independent of $\hat{\boldsymbol{\beta}}_1$.

For convenience of illustration, it can be shown that

$$\hat{\beta}_k | \mathbf{X}_D \sim N(\beta_k, V(\hat{\beta}_k)),$$

where

$$V(\hat{\beta}_k) = \frac{\sigma^2}{(1 - R_k^2) S_k^2},$$

R_k^2 is the coefficient of determination or R^2 in the regression of X_k on all other variables $(X_1, \ldots, X_{k-1}, X_{k+1}, \ldots, X_p)$, and $S_k^2 = \sum_{i=1}^{N}(X_{ik} - \overline{X}_k)^2$ is the corrected sum of squares with $\overline{X}_k = \sum_{i=1}^{N} X_{ik}/N$ for $k = 1, \ldots, p$. The corresponding test for the hypothesis H_0: $\beta_k = 0$ versus H_1: $\beta_k \neq 0$ is based on

$$t_k = \frac{\hat{\beta}_k}{\{\hat{V}(\hat{\beta}_k)\}^{1/2}}, \tag{A2}$$

where

$$\hat{V}(\hat{\beta}_k) = \frac{\hat{\sigma}^2}{(1 - R_k^2)S_k^2}. \tag{A3}$$

If the null hypothesis H_0: $\beta_k = 0$ is true, the statistic t_k is distributed as $t(N - p - 1)$, a central t distribution with $N - p - 1$ degrees of freedom, and H_0 is rejected at the significance level α if $|t_{XZ}| > t_{N-p-1,\alpha/2}$, where $t_{N-p-1,\alpha/2}$ is the upper $100(\alpha/2)$th percentile of the t distribution $t(N - p - 1)$.

Note that variance inflation factor (VIF) is a formal measure for identifying the extent of multicollinearity. In this case, the VIF of X_k is

$$VIF(X_k) = \frac{1}{1 - R_k^2}. \tag{A4}$$

For example, see Kutner, Nachtsheim, and Neter (2004, Sec. 10.5). With the definition in Equation (A4), the variance of $\hat{\beta}_k$ is directly linked to the widely used multicollinearity diagnostic of VIF through

$$V(\hat{\beta}_k) = \frac{\sigma^2 \cdot VIF(X_k)}{S_k^2}.$$

Also, the corresponding estimated variance in Equation (A3) can be rewritten as

$$\hat{V}(\hat{\beta}_k) = \frac{\hat{\sigma}^2 \cdot VIF(X_k)}{S_k^2}. \tag{A5}$$

When X_k has a substantial multicollinearity with the other predictor variables so that R_k^2 is substantially larger than 0, then $VIF(X_k)$ and $\hat{V}(\hat{\beta}_k)$ in Equations (A4) and (A5) are considerably inflated and even unbounded. The immediate and adverse consequence of large $\hat{V}(\hat{\beta}_k)$ is the t_k test in Equation (A2) may lead to false null hypotheses of no effect that disagree with prior knowledge and theoretical grounding. Another widely used multicollinearity diagnostic is the regressor correlation matrix determinant $|\mathbf{R}|$, where $\mathbf{R} = \mathbf{D}^{-1/2}(\mathbf{X}_C^{\mathrm{T}}\mathbf{X}_C)\mathbf{D}^{-1/2}$ is the regressor correlation matrix with diagonal matrix $\mathbf{D} = \mathrm{diag}(S_1^2, S_2^2, \ldots, S_p^2)$. The diagnostic of regressor correlation matrix determinant ranges from 0 when there is perfect multicollinearity to 1 when there is no multicollinearity.

Moreover, the resulting power function for the test H_0: $\beta_k = 0$ versus H_1: $\beta_k \neq 0$ is

$$P\{|t_k| > t_{N-p-1,\alpha/2}|\mathbf{X}_D\} = P\{|t(N-p-1,\Lambda)| > t_{N-p-1,\alpha/2}|\mathbf{X}_D\}, \quad \text{(A6)}$$

where $t(N-p-1,\Lambda)$ is the noncentral t distribution with $N-p-1$ degrees of freedom and noncentrality parameter

$$\Lambda = \frac{\beta_k}{\{V(\hat{\beta}_k)\}^{1/2}}.$$

Traditionally, the multiple regression model defined earlier is referred to as a fixed (conditional) model. The corresponding results would be specific to the particular values of the predictor variables that are observed or preset by the researcher. Under the random regression setup, the predictor variables $\mathbf{X}_i$, $i = 1,\ldots,N$, are assumed to have a joint probability density function $f(X_{i1},\ldots,X_{ip})$ and the form of $f(X_{i1},\ldots,X_{ip})$ does not depend on any of the unknown parameters $(\beta_0,\beta_1,\ldots,\beta_p)$ and σ^2. It is conceivable that the extended consideration of random feature associated with predictors complicates the fundamental statistical properties of the inferential procedures. However, the estimates of parameters and tests of hypotheses are the same under both fixed and random formulations. Nonetheless, the distinction between the two modeling approaches becomes important when unconditional or overall properties are to be evaluated.

Note that the observed values of $\mathbf{X}_i$, $i = 1,\ldots,N$, only represent one realization over the whole domain of $(X_1,\ldots,X_p)$. Interestingly, the unconditional mean $E[\hat{\beta}_k]$ of $\hat{\beta}_k$ remains unbiased because

$$E[\hat{\beta}_k] = E_X[E_Y\{\hat{\beta}_k\}] = E_X[\beta_k] = \beta_k,$$

where the expectations $E_Y[\cdot]$ and $E_X[\cdot]$ are taken with respect to the *iid* probability distributions $f(Y_i)$ and $f(X_{i1},\ldots,X_{ip})$ of Y_i and $\mathbf{X}_i = (X_{i1},\ldots,X_{ip})^{\mathrm{T}}$, respectively, $i = 1,\ldots,N$. Also, the unconditional variance $\nu(\hat{\beta}_k)$ of $\hat{\beta}_k$ is given by

$$\nu(\hat{\beta}_k) = E[(\hat{\beta}_k - \beta_k)^2] = E_X[E_Y\{(\hat{\beta}_k - \beta_k)^2\}] = \sigma^2 E_X\left[\frac{VIF(X_k)}{S_k^2}\right]. \quad \text{(A7)}$$

The power function in the context of random regression is defined as the expected value of the conditional power function given in Equation (A6) as follows:

$$\pi(t_k) = P\{|t_k| > t_{N-p-1,\alpha/2}\} = E_X[P\{|t(N-p-1,\Lambda)| > t_{N-p-1,\alpha/2}\}]. \quad \text{(A8)}$$

Likewise, the unconditional multicollinearity diagnostics of VIF and determinant of regressor correlation matrix are expressed as $\phi(X_k)$ and δ, respectively, where

$$\phi(X_k) = E_X[VIF(X_k)] \text{ for } k = 1, \ldots, p, \text{ and } \delta = E_X[|\mathbf{R}|]. \qquad \text{(A9)}$$

In general, there is no simple closed-form expression for the preceding quantities given in Equations (A7)–(A9) except in some special cases. Therefore, it is extremely cumbersome to evaluate the multidimensional integration with respect to the joint probability density distribution of $(X_1, \ldots, X_p)$. Instead, Monte Carlo integration provides a computationally feasible and practically accurate solution. Finally, the corresponding hypothesis testing procedures and power functions for the two one-sided tests of H_0: $\beta_k \leq 0$ versus H_1: $\beta_k > 0$ and H_0: $\beta_k \geq 0$ versus H_1: $\beta_k < 0$ and even nonzero minimum effect can be readily established but the details are not given here.