MEAN CENTERING IN MODERATED MULTIPLE REGRESSION: MUCH ADO ABOUT NOTHING

JEFFREY D. KROMREY AND LYNN FOSTER-JOHNSON
University of South Florida

Centering variables prior to the analysis of moderated multiple regression equations has been advocated for reasons both statistical (reduction of multicollinearity) and substantive (improved interpretation of the resulting regression equations). This article provides a comparison of centered and raw score analyses in least squares regression. The two methods are demonstrated to be equivalent, yielding identical hypothesis tests associated with the moderation effect and regression equations that are functionally equivalent.

The analysis of interaction effects between continuous variables in multiple regression has received considerable attention in recent years (e.g., Aiken & West, 1991; Jaccard & Wan, 1995; McClelland & Judd, 1993), although the methodology for such analysis has been available for at least 40 years (Saunders, 1955). An interaction effect indicates that the relation between a criterion variable and a predictor variable varies as a function of some third variable. This third variable is usually referred to as a *moderator* (Saunders, 1955). Moderator variables are common in behavioral research (Baron & Kenny, 1986). For example, Perlin, Menagham, Lieberman, and Mullen (1981) hypothesized moderating effects for both coping responses and social support on the relationship between stressful events and health. Similarly, Findley and Cooper (1983) hypothesized that the relationship between locus of control and academic achievement is moderated by demographic factors such as gender, race, and socioeconomic status.

An earlier version of this article was presented at the annual meeting of the Eastern Educational Research Association, Boston, 1996. Lynn Foster-Johnson is now a research statistician at the Amos Tuck Graduate School of Business, Dartmouth College, Hanover, NH.

Educational and Psychological Measurement, Vol. 58 No. 1, February 1998 42-67

The most common statistical test for interaction (or moderator) effects is accomplished with hierarchical multiple regression, in which differences in sample R^2 values between an additive model and a nonadditive model are tested (equivalently, for a single moderator component, the test of the regression weight for the product term may be used). Although alternative testing procedures have been recommended in the literature, such procedures subsequently have been shown to be incorrect (e.g., Cronbach, 1987; Dunlap & Kemery, 1987). As McClelland and Judd (1993) asserted, there has been "no credible published refutation of the appropriateness of [hierarchical multiple regression] as a test of moderator effects" (p. 377).

The frequently encountered difficulties in statistically detecting such effects in nonexperimental research have been attributed to a variety of factors, including measurement error (Dunlap & Kemery, 1988; Jaccard & Wan, 1995), multicollinearity (Morris, Sherman, & Mansfield, 1986), low residual variance of the product term in the regression equation (McClelland & Judd, 1993), residual variance heterogeneity (Alexander & DeShon, 1994), and even a natural consequence of multivariate normality (Fisicaro & Tisak, 1994). The degree to which each of these factors actually affects the detection of moderating variables remains unresolved. Consequently, the issues warrant consideration for their potential impact in detection of moderating variables.

Problems with multicollinearity in least squares regression are well documented, particularly with multiple regression models containing both main effects and interaction terms (see, for example, Cohen & Cohen, 1983). Although the least squares estimates of the regression coefficients remain unbiased, as multicollinearity increases, the determinants of the independent variables' covariance and correlation matrices approach zero and the standard errors of the coefficients increase. The resulting ill conditioning yields coefficients, *b*s, and an associated variance-covariance matrix (V_b) that are unstable. Small changes due to measurement or rounding error may be magnified, resulting in large changes in b and V_b. In addition, when multicollinearity is present, slight sampling fluctuations in the estimates of the covariances can result in great variability in the values and signs of least squares estimates of the coefficients. Finally, as a result of the increase in the expected distance between the vector of the least squares coefficients and the vector of true regression coefficients, estimates with excessively large values or unreasonable signs may result when extreme collinearity is present.

The method of centering the regressor variables by subtracting the sample mean from each observed value has been frequently recommended as a potential solution to problems in moderated multiple regression (MMR) analysis (see, for example, Cronbach, 1987; Smith & Sasaki, 1979; Yi, 1989). Mean centering has been offered as a simple data transformation that minimizes the multicollinearity in MMR. According to these authors, mean centering can increase the precision and stability of estimates by reducing the standard error, improve conditioning and stability of parts of the solutions, produce least squares estimates for the original model that can be calculated from those for modified variables, and yield coefficients for other variables in the model that are unaffected. Additional advantages suggested by some authors are that the regression weights obtained with centered data are more "meaningful" than those obtained with the uncentered data.

However, the degree to which centering actually affects the statistical tests for moderator effects is questionable. Cohen (1978) asserted that the tests for differences in R^2 between additive and nonadditive models are not affected by centering. The same conclusion was reached by Cronbach (1987). Although he advocated the use of centered regressors for interpretational reasons, Cronbach acknowledged that "only computational problems would cause analysis with the raw-score product to depart from the result by analysis with deviation scores" (p. 414). Furthermore, Mossholder, Kemery, and Bedeian (1990) argued that the interpretation of the regression weights is unchanged with centered data when the interpretation is conducted correctly. From this point of view, the arguments about the benefits of centering are specious and serve only to distract researchers from the multitude of pertinent issues in the conduct of inquiry.

The potential dilemma of "to center or not to center" emanates from three related issues: the substantive meaning of the elements in a regression equation, the interpretability of "main effects" in the presence of interaction effects, and the effects of multicollinearity on tests of hypotheses in linear models. Advocates of mean centering suggest that the degree of collinearity induced by the inclusion of the raw score product terms (or the inclusion of power terms in the case of polynomial regression) needed for moderated regression analysis adversely affects both the tests of hypotheses and the parameter estimates in the equations. Both effects may be avoided by centering the data prior to the computation of product or power terms. Additionally, according to some authors, the use of centered regressors allows one to clearly interpret main effects along with the "interactions" implied by the moderated regression equation.

The purpose of this article is to explicate the issues surrounding centering in nonadditive and nonlinear least squares regression models and to present the results of an empirical examination of the effects of centering in the analysis of data resulting from such models. Following a brief explication of centered and raw score regression models, the interpretation of the results of such models and the fundamental equivalence of these models will be illustrated using a heuristic data set. Second, the equivalence of centered and raw score regression analyses will be demonstrated using the results of a Monte Carlo study that investigated the effects of centering on the hypothesis tests associated with MMR. Finally, the issue of interpreting main effects in the presence of interactions in MMR will be addressed.

Centered and Raw Score Regression Models

This brief explication of centered and raw score regression models focuses on moderated regression analysis, in which one regressor is hypothesized to "moderate" or alter the relationship between another regressor and the criterion variable. In the interest of saving space, the application of the centering issue to nonlinear models (i.e., models of the form $Y = b_0 + b_1X_1 + b_2X_1^2 + e$) is not provided. However, the effects of power terms on both the correlation matrices and the results of regression analyses parallel those of product terms described here.

The simplest moderated regression model includes, as regressors, two variables (X_1 and X_2), together with a vector consisting of the products of the variables (X_1X_2):

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_1X_2 + e, \tag{1}$$

where the b_i are parameter estimates obtained with ordinary least squares and e is the residual of the regression.

A test of the null hypothesis that no moderation effect is present in the population from which the sample was obtained is provided by testing the statistical significance of b_3. Equivalently, but perhaps more informatively, this null hypothesis can be thought of as the difference in R^2 values between Equation 1 and the regression equation without the product term:

$$Y = b_0 + b_1X_1 + b_2X_2 + e. \tag{2}$$

If the R^2 in Equation 1 is statistically significantly larger than the R^2 value obtained from Equation 2, evidence of a moderator effect may be claimed.

If the two X variables are linearly transformed by subtracting the sample mean from each observed score prior to computing the product term and prior to estimating the regression parameters, the centered regression equations are obtained (with a lowercase x used to represent each regressor variable expressed as deviations from the sample mean, $x_i = X_i - \overline{X}_i$):

$$Y = b_0 + b_1x_1 + b_2x2 + b_3x_1x_2 + e \tag{3}$$

$$Y = b_0 + b_1x_1 + b_2x2 + e. \tag{4}$$

The algebra of least squares regression analysis provides the following results, which are apparently the source of much confusion in the literature:

1. The coefficient of determination for Equation 1 is equal to that of Equation 3.
2. The coefficient of determination for Equation 2 is equal to that of Equation 4.
3. The regression weights for X_1 in Equation 2 and x_1 in Equation 4 are equal to each other, as are the regression weights for X_2 in Equation 2 and x_2 in Equation 4.
4. The regression weight for the product term (b_3) in Equation 1 is equal to that for the product term in Equation 3.

5. The regression weights for X_1 in Equation 1 and x_1 in Equation 3 are not equal to each other, nor are the regression weights for X_2 in Equation 1 and x_2 in Equation 3.
6. The regression weights for x_1 and x_2 in Equation 3 are usually similar to those obtained in Equation 4, but the weights for X_1 and X_2 in Equation 1 are usually different from those obtained in Equation 2.

Interpreting Centered and Raw Score Regression Models

The effects noted above and their implications for the interpretation of moderated regression models will be illustrated with the use of a set of data constructed for illustrative purposes. Table 1 presents 10 observations with scores on measures of environmental stress level (X_1), quality of social support systems (X_2), and manifest level of stress symptoms (Y). Both raw scores and mean-centered values are provided. The centered values were obtained by subtracting the sample mean from each observed score on a variable. This, in effect, redistributes the scores on each variable from means of 39.8 and 29.7 (for X_1 and X_2, respectively) to means of zero on each variable. The cross-product terms, which are used to represent interaction in regression models, were obtained by multiplying X_1 and X_2 (or x_1 and x_2).

The substantive moderator variable hypothesis is that the quality of the social support system (the moderator variable) alters the relationship between environmental stress and the level of stress symptoms experienced. A statistically equivalent hypothesis is that the level of environmental stress (as a moderator variable) alters the relationship between quality of social support and manifest level of stress symptoms.

Correlations are provided in Table 2, with correlations between the raw score values above the main diagonal and correlations between the centered values below. Note that the correlations among the individual variables are identical whether the data are centered or not. A relatively low level of correlation is apparent between the level of environmental stress and the quality of the social support system (r = .18). Each of these variables is positively correlated with the criterion, manifest level of stress symptoms (r = .88 and r = .46, for level of environmental stress and quality of social support, respectively). The correlations with the cross-product term, however, show distinct differences with centering. In the raw data, the product term is highly correlated with both X_1 (r = .89) and Y (r = .94) and is moderately correlated with X_2 (r = .58). After centering the data, however, the product term (x_1x_2) has a small, negative correlation with x_1 (r = –.16) and nearly zero correlations with both x_2 (r = .08) and Y (r = .04).

Finally, the results of the regression analyses (Equations 1 through 4) are presented in Table 3. In the nonmoderated, additive models (Equations 2 and 4), the regression weights for environmental stress (X_1) and social support

Table 1
Hypothetical Data Set

	Manifest Symptoms Y	Raw Data			Centered Data		
		Environmental Stress X_1	Social Support X_2	Product Term X_1X_2	Environmental Stress x_1	Social Support x_2	Product Term x_1x_2
	16	20	20	400	−19.8	−9.7	192.06
	48	25	20	500	−14.8	−9.7	143.56
	80	50	40	2,000	10.2	10.3	105.06
	80	52	35	1,820	12.2	5.3	64.66
	120	57	39	2,223	17.2	9.3	159.96
	104	70	26	1,820	30.2	−3.7	−111.74
	64	60	22	1,320	20.2	−7.7	−155.54
	24	10	34	340	−29.8	4.3	−128.14
	40	24	31	744	−15.8	1.3	−20.54
	56	30	30	900	−9.8	0.3	−2.94
M	63.2	39.8	29.7	1,207.0	0.0	0.0	24.6
SD	33.4	20.3	7.4	717.7	20.3	7.4	127.2

Table 2
Correlations With Raw Score and Centered Data

	Manifest Symptoms Y	Environmental Stress X_1	Social Support X_2	Product Term X_1X_2
Y	1.00000	0.87879	0.46210	0.93709
X_1	0.87879	1.00000	0.18045	0.88735
X_2	0.46210	0.18045	1.00000	0.58012
X_1X_2	0.03991	−0.15732	0.08059	1.00000

Note. Raw score correlations are above the main diagonal and centered score correlations are below.

(X_2) are identical, whether the data are centered or not (columns 2 and 4 in Table 3), and identical values are obtained for the coefficient of determination ($R^2 = .87$). The only difference in the equations is in the intercept ($b_0 = -32.28$ for the raw data, and $b_0 = 63.20$ for the centered data).

With the moderated regression models (Equations 1 and 3), different regression weights are obtained for X_1 and X_2, but the weights for the product term (X_1X_2 and x_1x_2) are identical ($b_3 = .04$). As with the additive models, both sets of data provide the same value for the sample coefficient of determination ($R^2 = .89$).

Differences between the weights obtained in the additive and moderated models are substantial when raw scores are used but are small when centered data are used. With the centered data, the regression weights for x_1 and x_2 are

Table 3
Regression Analyses With Raw Scores and Centered Data

		Raw Score Regression Model		Centered Regression Model	
Variable	Regression Coefficient	Moderated (1)	Additive (2)	Moderated (3)	Additive (4)
Intercept	b_0	13.640055	−32.279258	62.233052	63.200000
Environmental stress (X_1)	b_1	0.230132	1.351973	1.395650	1.351973
Social support (X_2)	b_2	−0.234137	1.403055	1.327736	1.403055
Product (X_1X_2)	b_3	0.039243	—	0.039243	—
Coefficient of determination	R_2	0.8890	0.8675	0.8890	0.8675

nearly identical in the moderated and additive models (1.40 vs. 1.35 and 1.33 vs. 1.40 for x_1 and x_2, respectively). In contrast, the corresponding weights for the raw data are substantially different in the additive and moderated equations (0.23 vs. 1.35 and −0.23 vs. 1.40 for X_1 and X_2, respectively).

Overall, for the additive model, most of the elements in centered and raw score regression are identical (bivariate correlations, regression weights, and R^2 value). The only difference in the two analyses is the intercept. In any regression equation, the intercept is equal to the expected value of the criterion variable when all regressors are equal to zero. In the raw score equation, the "score" of zero for both level of environmental stress (X_1) and quality of social support systems (X_2) means a zero value on the instruments used to measure these variables, whether such a score is either conceptually or operationally possible. Thus, if it were possible for a participant to receive scores of zero on both of these scales, the expected level of manifest stress symptoms would be −32.28.

In contrast, in the centered equation, a score of zero for x_1 and x_2 is the sample mean for each of these variables. In such an observation, the expected value of the criterion variable is the sample mean of the criterion variable (63.20 in Table 1). Transformation of the individual raw scores by centering has transferred the "zero" point from an arbitrary zero to the sample mean of each variable. The individual participant scores on environmental stress and social support have been reexpressed in relation to the sample mean rather than an arbitrary value of zero on the two scales. Thus, the zero points in the centered data represent different values of the variables from the zero points in the raw data.

Despite this obvious difference between the centered and raw score data, the substantive results of the two regression equations are identical, and both equations yield the same predicted values for the same scores. For example, if the sample mean values of X_1 and X_2 (39.80 and 29.70 from Table 1) are used in Equation 2 (using the weights and intercept from column 2 of Table

3), the resulting predicted value is identical to the intercept of the centered equation:

$$Y = -32.279258 + (1.351973 \times 39.80) + (1.403055 \times 29.70) = 63.20.$$

Similarly, if raw scores of zero on each variable are centered (–39.8 for x_1 and –29.7 for x_2) and are used with the intercept and weights from Equation 4 (column 4 of Table 3), the resulting predicted value is identical to the intercept of the raw score equation:

$$Y = 63.20 + (1.351973 \times -39.80) + (1.403055 \times -29.70) = -32.279258.$$

Both equations lead to the same predicted values for the same scores, the only difference being that the value of zero is defined differently when the data are centered. The same substantive conclusions should be derived from the regression analysis, whether or not the data have been centered.

In considering the moderated regression models (Equations 1 and 3), the general interpretation of such models is that the partial regression slope of Y on X_1 (or x_1) is different for different values of X_2 (or x_2). The slope is a conditional one rather than being a constant slope across values of the other regressor. When asked about the relationship between environmental stress (X_1) and manifest stress symptoms (Y) in the moderated regression equation, the prudent data analyst should say "It depends on the quality of social support" (X_2). The conditional nature of the partial regression slopes is easily seen if Equation 1 is rearranged:

$$Y = (b_0 + b_2X_2) + (\{b_1 + b_3X_2\}X_1) + e. \tag{5}$$

For any specified value of social support (X_2), the intercept of the functional equation relating environmental stress (X_1) to manifest symptoms (Y) is ($b_0 + b_2X_2$), and the regression slope for environmental stress is ($b_1 + b_3X_2$). Substituting the parameter estimates for Equation 1 from Table 3, and assuming a low level of environmental support of 20, the equation relating environmental stress (X_1) to manifest stress symptoms is (Y)

$$Y = 13.640055 + (-0.234137 \times 20) + \{0.230132 + (0.039243 \times 20)\}X_1$$

$$Y = 8.957315 + (1.014992)(X_1).$$

This equation tells us that in a low social support environment, for each unit increase in environmental stress, manifest stress symptoms are expected to increase by approximately 1 unit.

Similarly, for a high level of social support of 40, the equation is

$$Y = 13.640055 + (-0.234137 \times 40) + \{0.230132 + (0.039243 \times 40)\}X_1$$

$$Y = 4.274575 + (1.799852)(X_1).$$

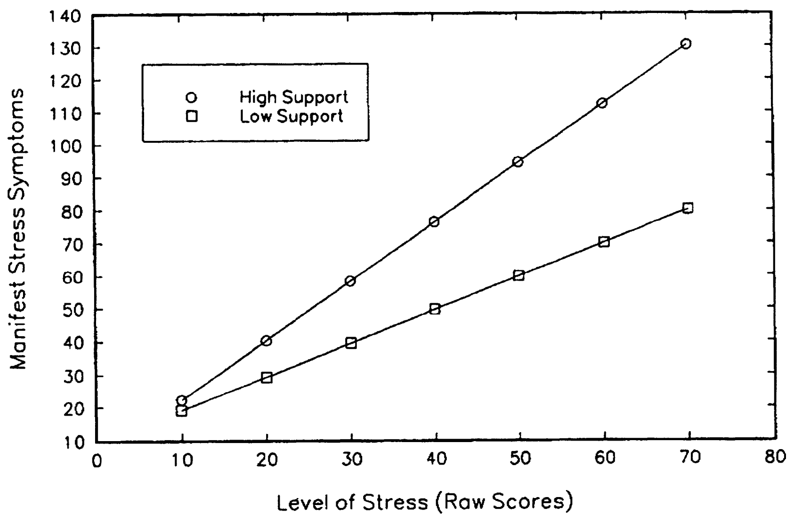

Figure 1. Regression lines for high and low support using raw data.

This equation tells us that in a high support environment, for each unit increase in environmental stress, manifest stress symptoms are expected to increase by approximately 1.8 units. These two equations are graphed in Figure 1.

Using a similar substitution of parameter estimates for Equation 3, from Table 3, and working with centered values of level of environmental stress (x_1) and quality of social support (x_2), the rearranged equation is

$$Y = (b_0 + b_2 x_2) + (\{b_1 + b_3 x_2\} x_1) + e. \quad (6)$$

A raw score of 20 on the quality of social support measure corresponds to a centered value of –9.7 (i.e., 20 – 29.7). Substituting this value in Equation 6 yields

$$Y = 62.233052 + (1.327736 \times -9.7) + \{1.395650 + (0.039243 \times -9.7)\} x_1$$

$$Y = 49.354013 + (1.014992)(x_1).$$

This regression equation tells us that in a low social support environment, for each unit increase in environmental stress, manifest stress symptoms are expected to increase by approximately 1 unit, exactly the same results as with the raw score equation.

Similarly, a raw score of 40 on the quality of social support measure is a centered value of 10.3. Substituting this value in Equation 6 yields

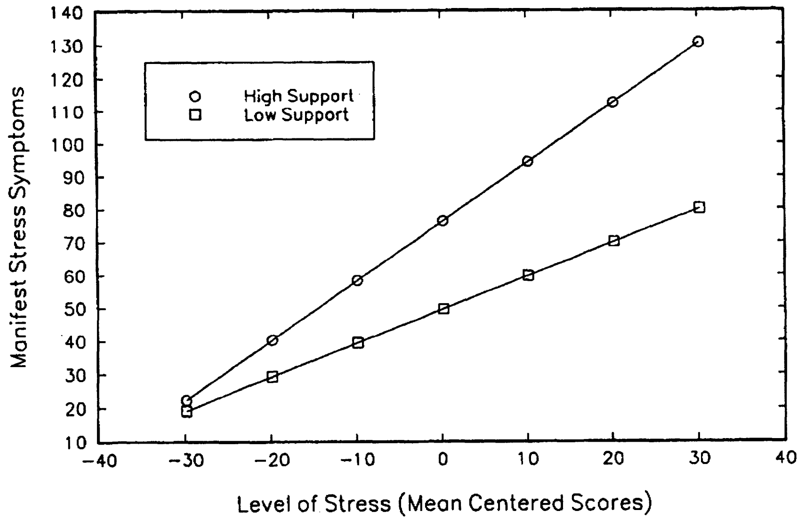

Figure 2. Regression lines for high and low support using mean-centered data.

$$Y = 62.233052 + (1.327736 \times 10.3) + \{1.395650 + (0.039243 \times 10.3)\}x_1$$

$$Y = 75.908732 + (1.799852)(x_1).$$

This equation, representing the high support environment, provides precisely the same regression slope as the raw score equation. That is, in a high support environment, for each unit increase in environmental stress, manifest stress symptoms are expected to increase by approximately 1.8 units. These two equations are graphed in Figure 2. The only difference in the graphs, and in the equations, is the scale of the abscissa. For the raw score equation, the scale value of zero represents a raw score of zero on the environmental stress measure. In contrast, for the centered equation, the scale value of zero represents a deviation score of zero or a raw score equal to the sample mean on the environmental stress measure. As with the additive regression models, the same substantive conclusions will be obtained in the interpretation of the moderated models, whether or not the data have been centered.

Effects of Centering on Hypothesis Tests in MMR

A second area in which centering is purported to improve MMR is that of the hypothesis tests for moderating effects. The reduction in collinearity afforded by centered data when product terms are included in correlation matrices should "logically" provide better (or at least different) hypothesis test results. An empirical comparison of such tests conducted with centered and raw scores, however, shows that these differences do not occur.

A Monte Carlo Study on Centered Versus Noncentered Models

To investigate the effects of centering on hypothesis testing for moderator effects, a Monte Carlo study was conducted that used regression models to generate data from populations evidencing (a) linear, (b) nonlinear, and (c) nonadditive relationships. In addition to the type of relationship characteristic of the population, five factors were manipulated in the study: (a) the correlation between the regressor variables, (b) the overall population R^2 of the regression model, (c) the effect size of the nonlinear or interaction terms (X_1X_1, X_2X_2, and X_1X_2), (d) the reliabilities of the regressors, and (e) sample size. Only models with two regressor variables were included in the study.

The magnitude of the correlation between the two regressors was controlled at levels ranging from .00 to .90. Four levels of overall R^2 of the population models were examined: .02, .13, .26, and .50. The first three levels represent small, medium, and large effect sizes for the population R^2, corresponding to f^2 values of .02, .15, and .35 (Cohen, 1988). The population R^2 value of .50 was included based on the review of correlational studies conducted by Jaccard and Wan (1995). In this review, the 75th percentile of the distribution of sample R^2s found in the psychological literature was .50. The magnitudes of the interaction component or the nonlinearity component were controlled at four levels, representing small, medium, and large effect sizes (Cohen, 1988), as well as a null condition.

Measurement error was simulated in the data (following the procedure used by Jaccard & Wan, 1995; and by Maxwell, Delaney, & Dill, 1984) by generating four normally distributed random variables for each observation (two to represent "true scores" on the regressors and two to represent errors of measurement). Fallible, observed scores on the regressors were calculated (under "classical" measurement theory) as the sum of the true and error variables. The reliabilities of the regressors were controlled by adjusting the error variances relative to the true score variances. Reliabilities were examined ranging from 0.40 to 1.00.

Sample sizes of 60, 175, and 400 were used. The larger two of these values represent the median and 75th percentile of sample sizes found in Jaccard and Wan's (1995) review of correlational studies in psychology. The small sample size (n = 60) was included to extend the results to small sample analyses. Five thousand samples of each size were generated for each condition in the Monte Carlo study. The use of 5,000 replications provides maximum 95% confidence intervals of ±.014 around the observed proportion of null hypotheses rejected (Robey & Barcikowski, 1992). Each sample was analyzed using uncentered data (generated with a population mean of 100 for each regressor) and with data centered around the sample mean for each regressor.

A procedure recommended by Lubinski and Humphreys (1990) was used to test for differences between linear, nonlinear, and nonadditive models. This procedure is a stepwise regression analysis in which increments in sample R^2 from moderated and nonlinear equations over that provided by a linear, additive equation are evaluated. The model that provides the largest increment in R^2 is selected as the best model, and the F test for change in R^2 is used to judge whether any model provides a statistically significantly better fit than the additive model. Although the Lubinski and Humphreys (1990) method for differentiating moderated and nonlinear models has been shown to be ineffective in many situations (for details, see Kromrey & Foster-Johnson, 1996), the hypothesis tests conducted as a part of the strategy provide a valid examination of possible differences between testing regression hypotheses using centered and uncentered data. For this study, a nominal alpha level of .05 was used for all hypothesis tests.

The Monte Carlo study was conducted using SAS, Versions 6.06 and 6.08. Normally distributed random numbers were generated using the RANNOR function, with a randomly selected seed value used for each sample generated. The components of the program were verified by comparing the results with the standard SAS output for benchmark data sets.

Results

In consideration of the comparison between raw data and mean-centered data, both approaches provided perfect agreement in the models selected using the Lubinski and Humphreys (1990) procedure. With 192 conditions examined in which an additive, linear model was the correct population model from which the data were generated, and 5,000 samples of each condition, a total of 960,000 samples were tested to compare the hypothesis tests between raw data and mean-centered data. In none of these samples was a difference detected.

The Type I error rates obtained using raw data and mean-centered data are presented in Table 4. The Lubinski and Humphreys (1990) procedure, being a stepwise regression analysis, provided poor control of Type I error in all conditions examined. In no condition was the empirical estimate of Type I error rate less than .075, and it was typically two to three times the nominal rate. In general, Type I error control was better with more highly correlated regressors and better with more reliable regressors.

Power estimates for the detection of moderated regression are provided in Tables 5 and 6 for the raw data and mean-centered data, respectively. Using the Lubinski and Humphreys (1990) procedures, these estimates represent

(Text continues on page 60)

Table 4
Empirical Type I Error Rates Using Lubinski and Humphreys's Procedure With Raw and Mean-Centered Data

			Correlation Between Regressors															
			$r12 = .00$				$r12 = .30$				$r12 = .60$				$r12 = .90$			
			Regressor Reliability				Regressor Reliability				Regressor Reliability				Regressor Reliability			
Effect Size	Additive Model R^2	Sample Size	0.40	0.60	0.80	1.00	0.40	0.60	0.80	1.00	0.40	0.60	0.80	1.00	0.40	0.60	0.80	1.00
Raw Data																		
0.00	.02	60	137	141	142	136	148	138	131	133	137	130	126	111	130	126	104	76
0.00	.02	175	145	139	144	146	141	133	138	134	145	124	127	117	135	116	104	85
0.00	.02	400	138	141	141	147	141	135	139	138	141	130	125	121	133	113	100	81
0.00	.13	60	144	143	138	143	143	138	129	133	136	134	123	111	131	118	104	80
0.00	.13	175	135	141	146	143	140	142	143	132	139	134	117	111	135	128	100	89
0.00	.13	400	143	140	144	144	145	148	135	131	139	122	120	116	132	118	104	87
0.00	.26	60	136	145	141	146	138	142	143	134	141	136	123	116	122	116	113	80
0.00	.26	175	147	137	140	140	139	140	123	139	134	132	128	106	132	114	105	76
0.00	.26	400	156	139	147	140	141	139	136	138	134	132	122	113	130	113	106	80
0.00	.50	60	147	146	134	141	134	137	133	129	135	126	117	118	131	118	94	81
0.00	.50	175	136	141	136	134	138	135	132	132	134	131	119	114	131	112	98	78
0.00	.50	400	146	141	140	141	136	148	135	143	136	127	123	112	141	118	102	85

Mean-Centered Data																			
0.00	.02	60	137	141	142	136	148	138	131	133	137	130	126	111	130	126	104	76
0.00	.02	175	145	139	144	146	141	133	138	134	145	124	127	117	135	116	104	85
0.00	.02	400	138	141	141	147	141	135	139	138	141	130	125	121	133	113	100	81
0.00	.13	60	144	143	138	143	143	138	129	133	136	134	123	111	131	118	104	80
0.00	.13	175	135	141	146	143	140	142	143	132	139	134	117	111	135	128	100	89
0.00	.13	400	143	140	144	144	145	148	135	131	139	122	120	116	132	118	104	87
0.00	.26	60	136	145	141	146	138	142	143	134	141	136	123	116	122	116	113	80
0.00	.26	175	147	137	140	140	139	140	123	139	134	132	128	106	132	114	105	76
0.00	.26	400	156	139	147	140	141	139	136	138	134	132	122	113	130	113	106	80
0.00	.50	60	147	146	134	141	134	137	133	129	135	126	117	118	131	118	94	81
0.00	.50	175	136	141	136	134	138	135	132	132	134	131	119	114	131	112	98	78
0.00	.50	400	146	141	140	141	136	148	135	143	136	127	123	112	141	118	102	85

Note. Estimates are based on 5,000 samples of each condition. Estimates have been multiplied by 1,000 and rounded.

Table 5

Empirical Power Estimates for Detecting Moderated Regression Component Using the Lubinski and Humphreys Procedure With Raw Data

			Correlation Between Regressors															
			$r12 = .00$				$r12 = .30$				$r12 = .60$				$r12 = .90$			
			Regressor Reliability				Regressor Reliability				Regressor Reliability				Regressor Reliability			
Effect Size	Additive Model R^2	Sample Size	0.40	0.60	0.80	1.00	0.40	0.60	0.80	1.00	0.40	0.60	0.80	1.00	0.40	0.60	0.80	1.00
0.02	.02	60	62	83	114	157	66	86	115	151	72	96	117	136	85	100	113	89
0.02	.02	175	101	167	269	399	107	179	288	406	135	220	304	393	164	250	309	263
0.02	.02	400	181	346	537	707	197	365	558	720	239	409	571	684	303	449	526	445
0.02	.13	60	61	81	104	148	61	86	111	157	69	89	119	137	82	99	109	93
0.02	.13	175	104	172	276	408	103	172	276	393	126	205	296	388	153	241	296	254
0.02	.13	400	164	330	533	707	184	357	555	726	231	394	573	700	303	459	505	446
0.02	.26	60	61	77	114	157	61	86	116	156	64	85	105	142	77	100	109	87
0.02	.26	175	98	153	261	400	97	168	271	401	117	200	292	396	153	239	287	259
0.02	.26	400	164	311	520	721	168	343	536	722	210	373	554	683	278	434	512	450
0.02	.50	60	55	72	102	151	54	71	105	155	61	82	112	136	67	89	101	96
0.02	.50	175	81	142	235	397	89	153	248	411	97	181	270	388	133	211	269	246
0.02	.50	400	120	265	475	720	148	299	500	711	183	345	524	686	240	395	486	453

0.15	.02	60	174	301	478	677	181	319	492	666	204	341	505	652	260	399	471	426
0.15	.02	175	401	686	846	899	411	679	832	897	435	644	810	893	493	657	749	712
0.15	.02	400	714	764	879	905	639	764	827	897	507	588	762	911	432	566	790	894
0.15	.13	60	149	300	488	677	168	317	488	668	200	340	517	633	265	366	456	433
0.15	.13	175	383	664	837	900	396	664	832	908	425	649	801	897	476	642	747	714
0.15	.13	400	676	865	886	904	622	756	837	906	518	597	762	906	445	568	798	876
0.15	.26	60	140	283	474	683	152	308	482	675	186	330	500	643	233	367	457	424
0.15	.26	175	359	652	837	903	367	662	827	901	408	640	797	894	476	646	744	710
0.15	.26	400	651	856	880	902	626	774	836	904	515	611	771	905	452	590	796	893
0.15	.50	60	114	239	419	685	125	253	445	688	161	287	462	650	202	338	443	427
0.15	.50	175	283	570	813	897	301	591	820	910	366	608	784	902	430	614	735	703
0.15	.50	400	539	830	895	899	554	776	847	902	529	639	785	909	498	618	810	881
0.35	.02	60	277	505	730	871	285	515	712	867	304	527	701	854	345	524	638	623
0.35	.02	175	621	826	852	900	568	736	823	909	488	584	746	907	452	586	783	882
0.35	.02	400	826	850	863	907	626	642	752	898	322	339	614	904	206	305	647	926
0.35	.13	60	261	480	717	877	271	504	720	874	290	526	703	843	342	511	629	635
0.35	.13	175	603	825	871	902	568	750	837	902	491	602	750	903	462	578	781	878
0.35	.13	400	813	854	866	901	641	667	757	906	342	365	628	912	225	326	671	920
0.35	.26	60	235	465	711	864	256	476	718	866	287	494	700	837	344	508	637	618
0.35	.26	175	580	815	870	910	540	740	828	895	487	609	764	906	477	593	784	878
0.35	.26	400	804	851	872	905	659	673	765	901	368	375	610	907	268	347	670	926
0.35	.50	60	192	407	677	873	216	435	691	866	253	466	676	852	304	497	627	630
0.35	.50	175	485	785	867	901	479	755	831	909	473	634	780	904	491	615	787	872
0.35	.50	400	763	856	860	900	658	702	782	904	436	440	648	902	317	397	698	927

Note. Estimates are based on 5,000 samples of each condition. Estimates have been multiplied by 1,000 and rounded.

Table 6
Empirical Power Estimates for Detecting Moderated Regression Component Using the Lubinski and Humphreys Procedure With Mean-Centered Data

			Correlation Between Regressors															
			$r12 = .00$				$r12 = .30$				$r12 = .60$				$r12 = .90$			
			Regressor Reliability				Regressor Reliability				Regressor Reliability				Regressor Reliability			
Effect Size	Additive Model R^2	Sample Size	0.40	0.60	0.80	1.00	0.40	0.60	0.80	1.00	0.40	0.60	0.80	1.00	0.40	0.60	0.80	1.00
0.02	.02	60	62	83	114	157	66	86	115	151	72	96	117	136	85	100	113	89
0.02	.02	175	101	167	269	399	107	179	288	406	135	220	304	393	164	250	309	263
0.02	.02	400	181	346	537	707	197	365	558	720	239	409	571	684	303	449	526	445
0.02	.13	60	61	81	104	148	61	86	111	157	69	89	119	137	82	99	109	93
0.02	.13	175	104	172	276	408	103	172	276	393	126	205	296	388	153	241	296	254
0.02	.13	400	164	330	533	707	184	357	555	726	231	394	573	700	303	459	505	446
0.02	.26	60	61	77	114	157	61	86	116	156	64	85	105	142	77	100	109	87
0.02	.26	175	98	153	261	400	97	168	271	401	117	200	292	396	153	239	287	259
0.02	.26	400	164	311	520	721	168	343	536	722	210	373	554	683	278	434	512	450
0.02	.50	60	55	72	102	151	54	71	105	155	61	82	112	136	67	89	101	96
0.02	.50	175	81	142	235	397	89	153	248	411	97	181	270	388	133	211	269	246
0.02	.50	400	120	265	475	720	148	299	500	711	183	345	524	686	240	395	486	453

0.15	.02	60	174	301	478	677	181	319	492	666	204	341	505	652	260	399	471	426
0.15	.02	175	401	686	846	899	411	679	832	897	435	644	810	893	493	657	749	712
0.15	.02	400	714	764	879	905	639	764	827	897	507	588	762	911	432	566	790	894
0.15	.13	60	149	300	488	677	168	317	488	668	200	340	517	633	265	366	456	433
0.15	.13	175	383	664	837	900	396	664	832	908	425	649	801	897	476	642	747	714
0.15	.13	400	676	865	886	904	622	756	837	906	518	597	762	906	445	568	798	876
0.15	.26	60	140	283	474	683	152	308	482	675	186	330	500	643	233	367	457	424
0.15	.26	175	359	652	837	903	367	662	827	901	408	640	797	894	476	646	744	710
0.15	.26	400	651	856	880	902	626	774	836	904	515	611	771	905	452	590	796	893
0.15	.50	60	114	239	419	685	125	253	445	688	161	287	462	650	202	338	443	427
0.15	.50	175	283	570	813	897	301	591	820	910	366	608	784	902	430	614	735	703
0.15	.50	400	539	830	895	899	554	776	847	902	529	639	785	909	498	618	810	881
0.35	.02	60	277	505	730	871	285	515	712	867	304	527	701	854	345	524	638	623
0.35	.02	175	621	826	852	900	568	736	823	909	488	584	746	907	452	586	783	882
0.35	.02	400	826	850	863	907	626	642	752	898	322	339	614	904	206	305	647	926
0.35	.13	60	261	480	717	877	271	504	720	874	290	526	703	843	342	511	629	635
0.35	.13	175	603	825	871	902	568	750	837	902	491	602	750	903	462	578	781	878
0.35	.13	400	813	854	866	901	641	667	757	906	342	365	628	912	225	326	671	920
0.35	.26	60	235	465	711	864	256	476	718	866	287	494	700	837	344	508	637	618
0.35	.26	175	580	815	870	910	540	740	828	895	487	609	764	906	477	593	784	878
0.35	.26	400	804	851	872	905	659	673	765	901	368	375	610	907	268	347	670	926
0.35	.50	60	192	407	677	873	216	435	691	866	253	466	676	852	304	497	627	630
0.35	.50	175	485	785	867	901	479	755	831	909	473	634	780	904	491	615	787	872
0.35	.50	400	763	856	860	900	658	702	782	904	436	440	648	902	317	397	698	927

Note. Estimates are based on 5,000 samples of each condition. Estimates have been multiplied by 1,000 and rounded.

not just power to detect a departure from a linear, additive model but also the power to accurately discriminate between the moderated model (with a product term) and a nonlinear model (with a power term).

A broad range of power estimates were obtained in the conditions examined in this study (ranging from .05 to .92). As expected, power increased with increasing effect size of the nonadditive term (providing mean power estimates of .26, .60, and .66 for the small, medium, and large effect sizes, respectively) and with increasing sample size (with mean power estimates of .35, .55, and .62 for sample sizes of 60, 175, and 400). Also, as expected, power increased with increasing reliability of the regressors. The average power with low reliability (r_{xx} = .40) was .32, whereas that of perfectly reliable regressors (r_{xx} = 1.00) was .68. Only a small but consistent decrease in power was observed as the correlation between regressors increased. The mean power estimates were .54, .53, .49, and .47 for values of r_{12} of .00, .30, .60, and .90, respectively. Finally, power was quite consistent across values of R^2 for the additive models (with mean power ranging only from .51 for R^2 = .02 to .49 for R^2 = .50).

Again, the interest for the current study is not in the success of the Lubinski and Humphreys (1990) procedure but in the differences in results between analyses based on raw data and those based on mean-centered data. For these models in which moderator effects were present in the populations from which samples were drawn, 576 conditions were examined with 5,000 samples of each condition. In more than 2.5 million samples, no differences were observed between the hypothesis tests conducted with centered data and those conducted with raw data.

Although the Lubinski and Humphreys (1990) procedure was not effective in controlling Type I error under any of the conditions examined, and although the procedure frequently showed very poor statistical power in identifying moderator effects, the effects of mean centering on the analyses were precisely nil. Whether the data were centered had no effect on any of the tests, and both analyses led to the same regression model for every sample examined.

Interpreting Main Effects in the Presence of Interactions

The interpretability of main effects in the presence of interactions is the third area in which centered models are purported to offer advantages relative to raw score models. Finney, Mitchell, Cronkite, and Moos (1984) described three perspectives on this issue, each of which provides a different definition of *main effect*. In the first approach, called the zero-point method, the main effect of X_1 is defined as the slope of the regression of Y on X_1 when X_2 is equal to zero. If the regressor variables are measured on a ratio scale, in which the value of zero represents the complete absence of the variable, this approach provides an estimate of one variable's relationship to the criterion variable when the other variable is completely absent. Lacking ratio scales, this method provides an arbitrary point (i.e., the point at which the value of zero is defined) at which the partial regression slope is assessed and called the main effect. The second approach, called the average effect approach, defines the main effect of X_1 as the average of the conditional slopes of X_1 in the sample. This is analogous to the definition of main effect in fixed-effects ANOVA models. The final approach, called the uninterpretable approach, is the assertion that main effects cannot be interpreted in the presence of interactions. This is the position taken by Cohen (1978), who indicated that although the partialed term $X_1X_2 \bullet X_1, X_2$ (representing the interaction in an MMR equation) is the interaction component, the partialed term $X_1 \cdot X_2, X_1X_2$ (representing the main effect of X1 in an MMR equation) is "arbitrary nonsense" (p. 861).

Obviously, the first approach is obtained with the regression weight for X_1 using raw score data, and the second approach is obtained with the regression weight using centered data. Further, either main effect may be obtained from either set of data by inserting the appropriate values of X_2 in the equations. The third approach asserts that attempting to interpret main effects in the presence of interactions is foolhardy. In general, the latter stance is probably the most reasonable, given that a comprehensive examination of the "simple effects" will provide the majority of the substantive interpretations of a moderated regression analysis.

To introduce the substantive issues involved in the interpretation of main effects, a fixed-effects factorial ANOVA model, in which both a main effect and an interaction effect are present, will be used. Presented in Table 7 are cell and marginal means from such a model. The data represent the effectiveness of three types of counseling on male and female clients. We can assume that the sample size was large enough to reject all of the pertinent null hypotheses (main effect for type of counseling, interaction effect, and subsequent contrasts). A graph of the cell means (the so-called simple effects) is provided in Figure 3.

In these data, the main effect of type of counseling is statistically significant (i.e., the marginal means for each counseling type are significantly different from each other), but the substantive nature of the effect is a "deceptive" one. Although the marginal mean for counseling Type A (53) is significantly higher than the other two marginal means (50 and 47), the substantive interpretation should be that the effectiveness of counseling depends on the gender of the participant. For female participants, counseling Type B was significantly more effective than either A or C, but for male participants counseling Type A was the most effective. Although Rosnow and Rosenthal have pointed out more than once (Rosnow & Rosenthal, 1989, 1995) that these simple effects are not the same as the statistical interaction component obtained by adjusting out the marginal means (the "pristine" interaction effects), the substantive interpretation and practical applications

Table 7
Fixed-Effects Factorial ANOVA With Deceptive Main Effect

	Type of Counseling			
Gender	A	B	C	Mean
Male	59	44	47	50
Female	47	56	47	50
Mean	53	50	47	

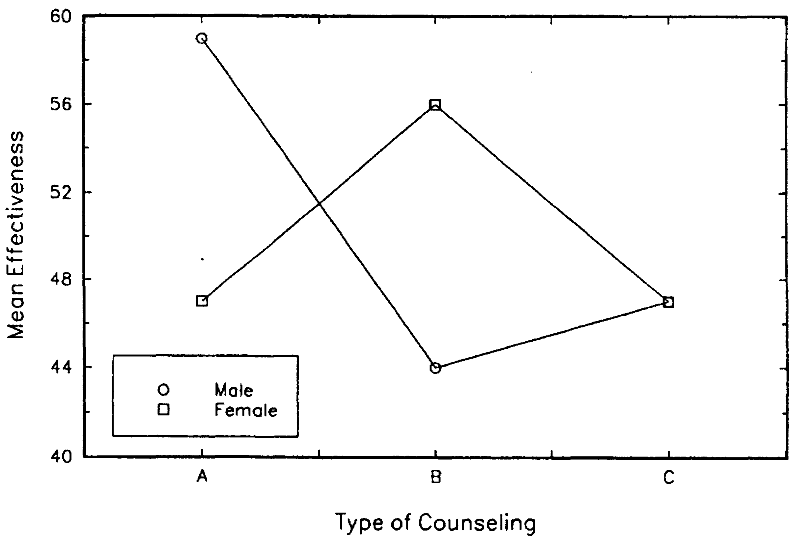

Figure 3. Fixed-effects ANOVA with deceptive main effect.

drawn from these data should be derived from the simple effects (Meyer, 1991) rather than from the interaction component in isolation.

A second set of hypothetical cell and marginal means is presented in Table 8 and graphed in Figure 4. Again, assume a large enough sample size to reject the relevant null hypotheses. For these data, the main effect for type of counseling is not deceptive. Counseling Type C provides better results than either A or B for both males and females. The interaction effect in these data relates to the relative effectiveness of counseling Types A and B for males and females. Although we need to conclude that the relative effectiveness of A and B depends on the client's gender, the superiority of C is unambiguous.

In MMR, the general notion of main effect is more poorly defined than in fixed-effects ANOVA models. However, either deceptive or nondeceptive

Table 8
Fixed-Effects Factorial ANOVA Without Deceptive Main Effect

	Type of Counseling			
Gender	A	B	C	Mean
Male	48	44	58	50
Female	44	48	58	50
Mean	46	46	58	

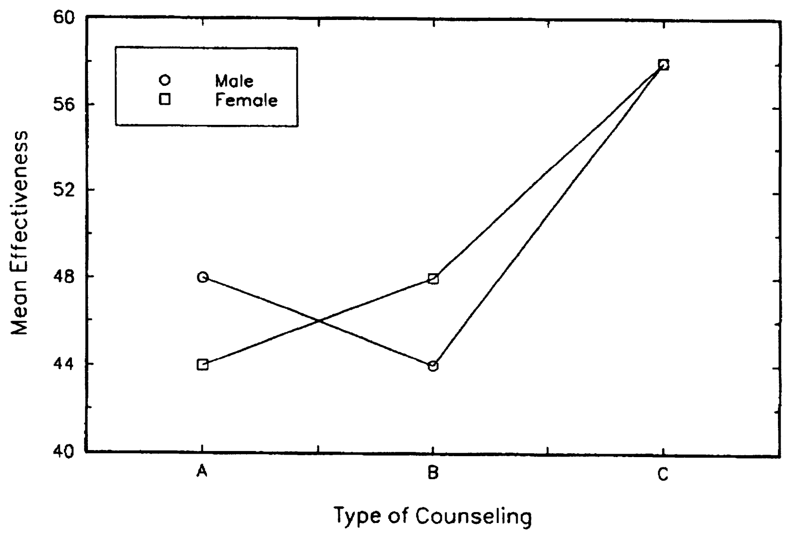

Figure 4. Fixed-effects ANOVA without deceptive main effect.

main effects may be present in data analyzed using MMR. For a main effect to be nondeceptive, the conditional slope of the regression line must be of the same algebraic sign and of a magnitude "large enough to matter" across the range of values of interest. Note that this circumstance is not readily discernible from the regression weights in either centered or raw score models. However, the extent of deceptiveness can be determined from either approach and can be readily communicated using a graphical method recommended by Tate (1984). For the presentation and interpretation of simple main effects in MMR, this method is elegant in its simplicity. The method provides a graph of the conditional slope of Y on X_1 as a function of X_2 (that is, the component $\{b_1 + b_3X_2\}$ in Equation 5). Such graphs obtained from the heuristic data set used earlier are provided in Figures 5 and 6 for the raw score

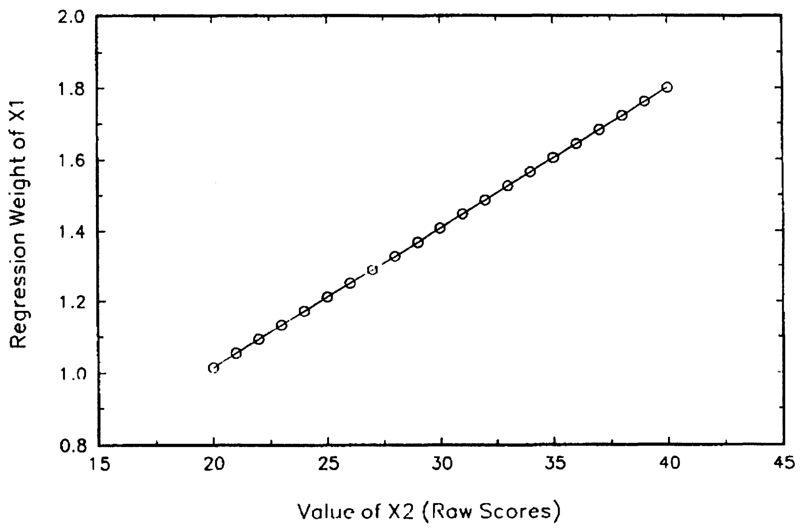

Figure 5. Regression coefficient of X_1 as a function of X_2 using raw data.

and centered models, respectively. Note that the slope of the line on these graphs is equal to b_3 (the regression weight for the product term in the equation). As with the other comparisons between raw score and centered regressions, these graphs are identical except for the scale of the abscissa.

If the main effect is deceptive, then clearly the effect should not be interpreted. The substantive meaning of the data analysis is contained in the conditional nature of the partial regression coefficients. If the main effect is not deceptive, cautious interpretation may be justified. However, two fundamental issues differentiate MMR from fixed-effects ANOVA models. First, most regression applications involve random variables as regressors rather than fixed levels of independent variables. In the heuristic example, researchers are probably interested in people with scores of 26 (for example) on the environmental stress scale and will generalize their results to these people even though this level of environmental stress was not observed in this sample. Second, we must remind ourselves not to extrapolate regression results beyond the range of values observed. If very high or very low values of either regressor were not included in the sample, the regression equation derived from the observed data may fit these unobserved values very poorly.

In general, the notion of main effect in the presence of an interaction effect is a more ambiguous construct in MMR than in fixed-effects ANOVA models. Although a part of the relationship between X_1 and Y that is statistically independent of both X_2 and X_1X_2 is calculated in conducting MMR, whether or not the data are centered, researchers must be cognizant of the meaning (if

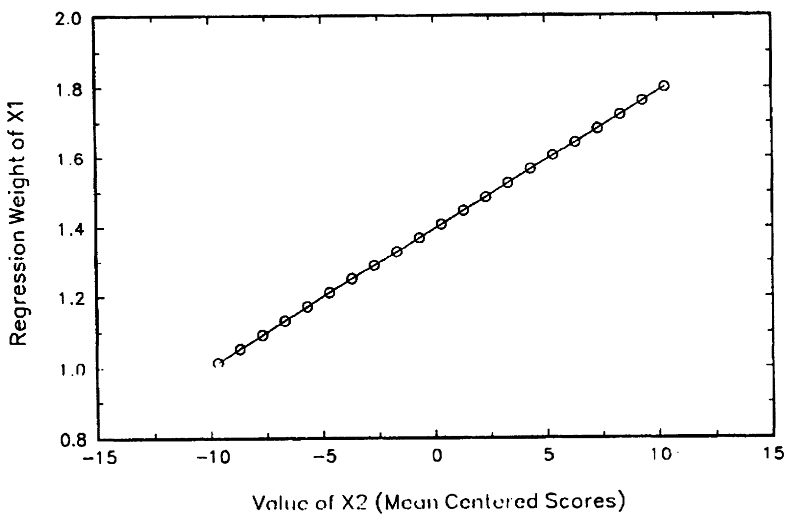

Figure 6. Regression coefficient of X_1 as a function of X_2 using mean-centered data.

any) that is represented by this component. Obviously, an attempt to interpret "arbitrary nonsense" (Cohen, 1978) is worse than a waste of time—it is dissemination of misinformation. The important issue is for the data analyst to think about the substantive meaning of the entire regression equation, whether centered or raw score, and to communicate this meaning to the reader. If such communication is improved by using the partial regression slope at the value of zero for the moderator variable, then the regression weight from the raw score equation should be used. Conversely, if the substantive meaning of the regression can be conveyed by using the average of the sample's partial regression slopes, then these should certainly be calculated and reported. However, to equate either of these numbers to the notion of a main effect in fixed-effects factorial ANOVA models invites much confusion and potential misinterpretation.

Conclusion

As the social sciences have developed, many relationships between variables have been hypothesized that go beyond simple, additive, linear relationships (Cortina, 1993). Procedures for analyzing data to empirically test such theories are of importance because of the pivotal role such analyses play in the interpretation of research. It is important that methodologists direct their attention to issues in data analysis that may provide improvements in field research and analysis techniques. Little is gained, and much energy is lost, by devoting attention to methodological choices that appear different on the surface but are actually identical. Such is the case with mean centering in least squares regression analysis.

We have attempted in this article to demonstrate that the equations obtained with centered and raw data are equivalent, that the results of hypothesis testing with either type of data are exactly the same, and that neither approach provides a viable vehicle for the interpretation of main effects in MMR. Although centering has important implications for both interpretations and hypothesis testing in hierarchical models (see, for example, Kreft, de Leeuw, & Aiken, 1995), with least squares regression analysis, the conclusion reached by Cohen (1978) is still the best—"one might just as well not bother" (p. 865).

References

Aiken, L., & West, S. (1991). *Multiple regression.* Newbury Park, CA: Sage.

Alexander, R. A., & DeShon, R. P. (1994). Effect of error variance heterogeneity on the power of tests for regression slope differences. *Psychological Bulletin, 115,* 308-314.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182.

Cohen, J. (1978). Partialed products are interactions; partialed powers are curve components. *Psychological Bulletin, 85,* 858-866.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cortina, J. M. (1993). Interaction, nonlinearity and multicollinearity: Implications for multiple regression. *Journal of Management, 19,* 915-922.

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

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.

Findley, M. J., & Cooper, H. M. (1983). Locus of control and academic achievement: A literature review. *Journal of Personality and Social Psychology, 44,* 419-427.

Finney, J. W., Mitchell, R. E., Cronkite, R. C., & Moos, R. H. (1984). Methodological issues in estimating main and interactive effects: Examples from coping/social support and stress field. *Journal of Health and Social Behavior, 25,* 85-98.

Fisicaro, S. A., & Tisak, J. (1994). A theoretical note on the stochastics of moderated multiple regression. *Educational and Psychological Measurement, 54,* 32-41.

Jaccard, J., & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin, 117,* 348-357.

Kreft, I.G.G., de Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research, 30,* 1-21.

Kromrey, J. D., & Foster-Johnson, L. (1996). *Statistically differentiating between interaction and nonlinearity in multiple regression analysis: A Monte Carlo investigation of a recommended strategy*. Paper presented at the annual meeting of the American Educational Research Association, New York City.

Lubinski, D., & Humphreys, L. G. (1990). Assessing spurious "moderator effects": Illustrated substantively with the hypothesized ("synergistic") relation between spatial and mathematical ability. *Psychological Bulletin, 107*, 385-393.

Maxwell, S. E., Delaney, H. D., & Dill, C. A. (1984). Another look at ANCOVA versus blocking. *Psychological Bulletin, 95*, 136-147.

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin, 114*, 376-389.

Meyer, D. L. (1991). Misinterpretation of interaction effects: A reply to Rosnow and Rosenthal. *Psychological Bulletin, 110*, 571-573.

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.

Mossholder, K. W., Kemery, E. R., & Bedeian, A. G. (1990). On using regression coefficients to interpret moderator effects. *Educational and Psychological Measurement, 50*, 255-263.

Perlin, L. I., Menagham, E. G., Lieberman, M. A., & Mullan, J. T. (1981). The stress process. *Journal of Health and Social Behavior, 22*, 337-356.

Robey, R. R., & Barcikowski, R. S. (1992). Type I error and the number of iterations in Monte Carlo studies of robustness. *British Journal of Mathematical and Statistical Psychology, 45*, 283-288.

Rosnow, R. L., & Rosenthal, R. (1989). Definition and interpretation of interaction effects. *Psychological Bulletin, 105*, 143-146.

Rosnow, R. L., & Rosenthal, R. (1995). Some things you learn aren't so: Cohen's paradox, Asch's paradigm, and the interpretation of interaction. *Psychological Science, 6*, 3-9.

Saunders, D. R. (1955). The "moderator variable" as a useful tool in prediction. In *Proceedings of the 1954 invitational conference on testing problems* (pp. 54-58). Princeton, NJ: Educational Testing Service.

Smith, K. W., & Sasaki, M. S. (1979). Decreasing multicollinearity: A method for models with multiplicative functions. *Sociological Methods and Research, 8*, 35-36.

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

Yi, Y. (1989). On the evaluation of main effects in multiplicative regression models. *Journal of the Marketing Research Society, 31*, 133-138.