The Detection and Interpretation of Interaction Effects Between Continuous Variables in Multiple Regression

James Jaccard, Choi K. Wan, and Robert Turrisi
University at Albany
State University of New York, Albany

Issues in the detection and interpretation of interaction effects between quantitative variables in multiple regression analysis are discussed. Recent articles by Cronbach (1987) and Dunlap and Kemery (1987) suggested the use of two transformations to reduce "problems" of multicollinearity. These transformations are discussed in the context of the conditional nature of multiple regression with product terms. It is argued that although additive transformations do not affect the overall test of statistical interaction, they do affect the interpretational value of regression coefficients. Factors other than multicollinearity that may account for failures to observe interaction effects are noted.

Morris, Sherman, and Mansfield (1986) have noted a persistent failure of psychologists to detect interaction effects between continuous variables in multiple regression analysis. They suggested that multicollinearity between product terms and their constituent parts may be the source of the problem and proposed a form of principal components analysis as a remedy. In two later articles, Cronbach (1987) and Dunlap and Kemery (1987) point out limitations with the approach of Morris et al. and specify alternative methods for coping with multicollinearity. Cronbach suggests that interaction effects be evaluated by the traditional hierarchical regression approach advocated by Cohen and Cohen (1983). However, he suggests that an additive transformation be performed on the predictor variables prior to the formation of the product term, and that the dependent variable (Y) be regressed onto the transformed variables and their product for purposes of evaluating the interaction effect. The transformation for a given predictor involves subtracting the mean of the predictor variable from each individual's raw score on that predictor, thus forming deviation scores. Cronbach suggests that such a transformation will yield low correlations between the product term and the component parts of the term. This is desired, he asserts, because it decreases the probability of computational errors and

Preparation of this article was supported by NICHD grant HD2215201 and NIAAA grant AA0687501. We would like to thank George Alliger, Jacob Cohen, William Dillon, and Scott Maxwell for their helpful comments on a previous draft of the manuscript. Requests for reprints should be sent to James Jaccard, Department of Psychology, State University of New York, Albany, New York 12222.
because “the bs are closely related to the simple $X$ and $Y$ effects and the interaction” (p. 416). Dunlap and Kemery examined the same issues as Cronbach, but suggested a transformation that involved standardizing the predictor variables. The results of the hierarchical test of the interaction effect are identical for both the Cronbach and Dunlap and Kemery transformations.

Based on these articles, a researcher might decide to use either transformation when testing for statistical interaction. The implication is that the choice of transformation is somewhat arbitrary, because the overall test of the interaction effect will yield identical results. Although this is substantively correct, we argue that the choice of a transformation has interpretative implications for the analysis and should not be arbitrary. Cronbach (1987) and Dunlap and Kemery (1987) also fail to provide insights into the problem motivating the original Morris et al. (1986) article, namely the persistent failure by psychologists using moderated regression analysis to observe significant interaction effects. The implication of the Cronbach article is that multicollinearity will only be a problem when it leads to computational errors within current computer algorithms. It is unlikely that the high degree of multicollinearity required for this to occur has been present in the majority of empirical evaluations of moderated multiple regression. Thus, other factors probably are operating that make it difficult to correctly detect moderated relationships. In this article, we will suggest what some of these others factors might be.

**Additive Transformations and Interpretive Issues**

Consider the case of three continuous variables, where the investigator is interested in the effects of two independent variables ($X1$ and $X2$) on a dependent variable ($Y$). The test of an additive (or main effects) model for predicting $Y$ from $X1$ and $X2$ typically takes the form of a least squares regression approach such that

$$Y = a + b1'X1 + b2'X2 + e,$$

where $a$ = the least squares estimate of the intercept, and $b1'$ and $b2'$ = the least squares estimates of the population regression coefficients for $X1$ and $X2$, respectively, and $e$ is a residual term. The sample multiple correlation coefficient, $R$, is an index of overall model fit (in the sample), and the regression coefficients represent estimates of the effects of an $X$ variable on $Y$, holding all other $X$ variables constant. For the case of interaction effects, a multiplicative term is formed, $X1X2$, which is said to encompass the interaction effect, yielding a three term equation:
J. Jaccard, C. Wan, and R. Turrisi

If an interaction effect is present, then the difference between the $R^2$ in Equations 1 and 2 should be statistically significant (barring a Type II error). The $F$ test in such a hierarchical regression strategy yields the same substantive result as a $t$-test of the $b_3$ coefficient for the multiplicative term.

The interpretation of regression coefficients in Equations 1 and 2 is distinct. In Equation 1, a regression coefficient estimates the effects of the independent variable on the dependent variable, across the levels of the other independent variable. That is, $b_1$ reflects the trends of change in $Y$ with changes in $X_1$ at various levels of $X_2$, and $b_2$ reflects the trends of change in $Y$ with changes in $X_2$ at various levels of $X_1$. In contrast, in Equation 2, the regression coefficients for $X_1$ and $X_2$ reflect conditional relationships: $b_1$ reflects the influence of $X_1$ on $Y$ when $X_2$ equals zero, and $b_2$ reflects the influence of $X_2$ on $Y$ when $X_1$ equals zero. The coefficient $b_3$ represents an interaction effect in that it estimates the change in the slope of $Y$ on $X_1$ given a one unit change in $X_2$ (or, alternatively, the change in the slope of $Y$ on $X_2$ given a one unit change in $X_1$, depending on how one conceptualizes the interaction).

The distinctions between the regression coefficients in Equations 1 and 2 hold with equal vigor for the standard errors associated with the coefficients: The standard errors for regression coefficients in Equation 1 reflect estimates of sampling error across levels of the independent variables. In contrast, the standard errors for regression coefficients in Equation 2 are conditional and reflect sampling error at particular levels of the independent variables. The standard error for $b_1$ in Equation 2 estimates sampling error for the regression coefficient when $X_2$ equals zero. Similarly, the standard error for $b_2$ in Equation 2 estimates sampling error for the regression coefficient when $X_1$ equals zero. The standard error for $b_3$ estimates sampling error for the effect of $X_1$ on $Y$ with a one unit change in $X_2$ (or, alternatively, the effect of $X_2$ on $Y$ with a one unit change in $X_1$).

Both the $b_1$ regression coefficient and its associated standard error will differ as a function of $X_2$, given the presence of statistical interaction. Although the derivation of $b_1$ and $b_2$ at selected values of $X_2$ and $X_1$, respectively, appear in several regression texts, the derivation of the corresponding standard errors is less well known. Focusing on Equation 2, if we factor $X_1$ from the terms $b_1X_1$ and $b_3X_1X_2$, we obtain the value of $b_1$ for any given value of $X_2$:

\[ Y = a + b_1X_1 + b_2X_2 + b_3X_1X_2 + e. \]

1 Technically, the models implied by Equations 1 and 2 are both conditional in nature. For Equation 1, the focus is on the conditional expectation of $E(Y|X_1,X_2)$, and for Equation 2, the focus is on the conditional expectation of $E(Y|X_1,X_2 = 0)$.

2 For pedagogical reasons, in the remainder of this article, we will interpret interaction effects as if $X_2$ is the moderator variable.
Similarly, the value of $b_2$ at any given value of $X_1$ is

$$b_2 \text{ at } X_1 = b_2 + (b_3)(X_1).$$

The corresponding standard error for $b_1$ at any given value of $X_2$ is

$$s(b_1 \text{ at } X_2) = [\text{var}(b_1) + X_2^2 \text{ var}(b_3) + 2X_2 \text{ cov}(b_1, b_3)]^{1/2},$$

and the standard error for $b_2$ at a particular value of $X_1$ is

$$s(b_2 \text{ at } X_1) = [\text{var}(b_2) + X_1^2 \text{ var}(b_3) + 2X_1 \text{ cov}(b_2, b_3)]^{1/2}.$$

Consider the following numerical example. For 125 hypothetical subjects, the intention to use birth control ($Y$) was predicted from the subjects’ attitude toward birth control ($X_1$) and the perceived normative pressures to use birth control ($X_2$). The $X_1$ and $X_2$ variables were measured on scales that ranged from 1 to 5, and the $Y$ variable was measured on a scale from 0 to 30. The $X_1$ and $X_2$ scores can be conceptualized in terms of a $5 \times 5$ factorial design, and the mean scores for each cell of the design are reported in Table 1. There are equal $n$ in each cell ($n = 5$). A small degree of within cell variability was introduced by allowing four of the five scores in a given cell to deviate one unit from the cell mean. Inspection of Table 1 shows an orderly, monotonic trend of changes in the slope of $Y$ on $X_1$ across the levels of $X_2$. When $X_2$ equals 1, the slope of $Y$ on $X_1$ is 1.00, and with every one unit that $X_2$ increases, the slope increases by 1.00 units.

The multiple $R$ for the two term additive model is 0.901 and the regression equation is

$$Y = -4.00 + 3.00 \cdot X_1 + 2.00 \cdot X_2 + e.$$ 

The multiple $R$ for the three term model is 0.968 and the regression equation is

$$Y = 5.00 + .00 \cdot X_1 + -1.00 \cdot X_2 + 1.00 \cdot X_1X_2 + e.$$ 

The coefficient $b_1 (.00)$ reflects the number of units that $Y$ is predicted to change given a one unit increase in $X_1$ when $X_2$ equals zero and the coefficient $b_2$ reflects the number of units that $Y$ is predicted to change given a one unit change in $X_2$ when $X_1$ equals zero. Note that the coefficient $b_3$ captures the changes in the slope of $Y$ on $X_1$ with changes in $X_2$ (i.e., for every one unit that $X_2$ changes, the slope of $Y$ on $X_1$ changes 1.00 units).
Table 1
Cell Means as a Function of X1 and X2

<table>
<thead>
<tr>
<th>X1</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>25</td>
<td>21</td>
<td>17</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>17</td>
<td>14</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>13</td>
<td>11</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Now consider the effects of an additive transformation on one of the X variables. Suppose we transform the normative pressure scores (X2) by subtracting the mean of X2 from each score (in this case, the mean is 3.00). Such a transformation is traditionally referred to as centering. This transformation will leave unchanged the values of b2 and b3, but will alter the values (and standard errors) of b1 and the intercept. The pre-transformation regression equation is

\[ Y = 5.00 + .00 X1 + -1.00 X2 + 1.00 X1X2 + e, \]

and the post-transformation regression equation is

\[ Y = 2.00 + 3.00 X1 + -1.00 X2 + 1.00 X1X2 + e. \]

The change in b1 in the two equations (from 0.00 to 3.00) occurs because the conditional relationship of the influence of X1 on Y is being evaluated at a different zero point than was originally the case. In the pre-transformation analysis, the zero point was based on a scale that was 3.00 units higher than the transformed X2 score, whereas after centering, the zero point occurs at the mean. In fact, if in the original analysis we wanted to evaluate the impact of X1 on Y at the mean of X2, we could do so by substituting the mean value for X2 in Equation 9 and then, by algebraic manipulation, calculate the slope of Y on X1:

\[ Y = 5.00 + .00 X1 + (-1.00)(3.00) + (1.00) X1 \cdot 3.00 = 2.00 + 3.00 X1. \]

Note that the observed slope equals the slope for X1 in Equation 10 where X2 has been centered. Application of Equation 5 to the above would yield the
identical standard error for $b_1$ as Equation 10. Centering has no effect on the substantive evaluation of the effect of $X_1$ on $Y$ at a given value of $X_2$. It only changes the value of $X_2$ being evaluated (because with centering, a zero on $X_2$ corresponds to the mean, whereas without centering, this is not necessarily the case).

It can be seen that Cronbach's (1987) contention that $b_1$ and $b_2$ using centered data "more closely relates to the simple $X$ and $Y$ effects" is somewhat misleading. With centered data, $b_1$ reflects the influence of $X_1$ on $Y$ at the mean $X_2$ score. Similarly, $b_2$ reflects the influence of $X_2$ on $Y$ at the mean $X_1$ score. Cronbach's transformation is desirable because it does indeed reduce potential problems with multicollinearity when testing for the presence of statistical interaction. However, one must go beyond concern with just this criterion and also consider the implications of a transformation for the meaningful interpretation of the nature of an interaction effect, not just the presence of that effect. In this regard, some transformations are more useful than others.

In traditional analysis of variance (ANOVA) paradigms, two methods of interaction decomposition are used, simple main effects analysis and interaction comparisons (see Keppel, 1982). There is a direct analog to both procedures in multiple regression with product terms. Simple main effects analysis focuses on the statistical significance of the effects of an independent variable ($X_1$) on a dependent variable ($Y$) at selected levels of a moderator variable ($X_2$). This corresponds to the specification of the slope of $Y$ on $X_1$ at theoretically or empirically meaningful values of $X_2$, and a corresponding test of significance of those effects. To conduct the analysis, one needs to specify values of $X_2$ where one wishes to evaluate the effects of $X_1$ on $Y$. In the absence of theory to guide this choice, a reasonable strategy is to evaluate the effects of $X_1$ on $Y$ at low, medium, and high values of $X_2$, where low might be defined as one standard deviation below the mean, medium as at the mean, and high as one standard deviation above the mean. This can be accomplished by centering data and then calculating the appropriate three term regression equation. Equations 3-6 are then used to define the relevant coefficients and their associated standard errors for low, medium and high scores. With centered data, a score of zero on the moderator variable corresponds to a medium value, a score of $(+1)(sd)$ corresponds to a high value and a score of $(-1)(sd)$ corresponds to a low value, where $sd = \text{the standard deviation of the moderator variable}$. A $t$-ratio to test the significance of $b_1$ at a given value of $X_2$ is formed by dividing a given coefficient by its standard error. Formulas for computing the standard errors from traditional computer output are provided in the appendix.

In contrast to simple main effects analysis, the approach of interaction comparisons focuses on the formal comparison of slope differences as one moves from one value of the moderator variable to another. This information is readily available in $b_3$ (and its associated test of significance): Again using
centered data, for every one unit that X2 changes, the slope of Y on X1 changes b3 units. Thus b3 contains the core information of an interaction comparison analysis.

Let us now consider the transformation proposed by Dunlap and Kemery (1987), which involves standardizing the Y and X variables in Equation 2 prior to forming the cross-product term. This transformation incorporates the centering process advocated by Cronbach (1987), but takes the additional step of dividing the centered variables by their respective standard deviations. The transformation, like Cronbach's transformation, eliminates threats to the test of the overall interaction effect due to multicollinearity. However, the transformation has direct implications for the interpretation of the regression coefficients which are distinct from that of centering. Specifically, the Dunlap and Kemery (DK) transformation yields a standardized solution, whereas the Cronbach transformation maintains an unstandardized format. All of the equations described above are applicable to DK-transformed data, but now the unstandardized regression coefficients are interpreted in the spirit of standardized coefficients (i.e., they refer to standard scores). To be sure, the DK analysis differs from traditional standardized regression analysis in that a non-zero intercept term may result, but this generally will be of little theoretical consequence. Let ZY represent the standard score equivalent of Y, Z1 represent the standard score equivalent of X1, and Z2 represent the standard score equivalent of X2. In DK transformed data, b1 will reflect the number of Z scores that Y changes given a one Z score change in X1 when Z2 = 0, that is, when X2 equals its mean value. b2 will reflect the number of Z scores that Y changes given a one Z score change in X2 when Z1 = 0, that is, when X1 equals its mean value. b3 indicates that for every one Z score that X2 increases, the slope of ZY on Z1 changes by b3 units. In practice, the values of b1 and b2 in the standardized version of Equation 2 will tend to approximate the values of the corresponding standardized regression coefficients in Equation 1.

We should note that the DK transformation requires that the standardization process occur prior to the formation of the product term. Traditional standardized solutions in which all variables (including the product term) are standardized is not appropriate for the meaningful interpretation of moderated multiple regression. The key to interaction analysis is the ability to factor X1 from the weighted product term b3X1X2 and b1X1 (see Equation 3). To do so with standardized variables, the terms in the equation need to equal b3Z1Z2 and b1Z1. This is not the case in traditional standardized solutions. Rather, the product term represents the standardization of X1X2 (or Z1Z2) rather than the product of the separate Z scores, yielding b3Z1Z2 and b1Z1. This undermines the conditional interpretation of b3.

There currently exists controversy on the use of standardized variables in regression analysis. Many social scientists prefer the use of standardized
measures because the variables are measured on a common metric (i.e., with a mean of zero and a standard deviation of 1). However, strong arguments have been presented against the use of such measures in certain situations (see, for example, Kim & Ferree, 1981), especially in the context of causal analysis via structural equation models. The major problem with standardized coefficients is that they lack the property of causal invariance. Causal relationships typically are conceptualized in terms of change. A variable \( X \) is said to be a cause of \( Y \) if changes in \( X \) produce changes in \( Y \). Because regression coefficients also focus on change, it is natural to study them in the context of causal models. If a causal relationship is identical in each of a set of groups, then the coefficients on which an analysis is performed should reflect this invariance. Situations arise where unstandardized regression coefficients properly reflect causal invariance, whereas standardized regression coefficients do not. Consider the following example. Assume in a population that \( Y \) is completely determined by \( X_1 \) and \( X_2 \), in accord with the following linear model:

\[
Y = 0.60 X_1 + 0.80 X_2.
\]

The value of the intercept is zero and there is no residual term. Assume also that the standard deviations of \( X_1 \) and \( X_2 \) are both 1.0. Now suppose that three subsets of scores are randomly selected from the population. Table 2 presents scores that might be observed. The scores in each group represent atypical, but nevertheless, plausible random samples. Note that each \( Y \) score is completely specified by Equation 11. If one computes unstandardized regression coefficients in each sub-group, the result will be the generating equation (i.e., \( Y = 0.60 X_1 + 0.80 X_2 \)). The coefficients are invariant. This is not true of the standardized coefficients. Specifically, the standardized coefficients for the three sub-groups would be 0.80, 0.56 and 0.94, respectively. These coefficients are not invariant across sub-groups (or over replications of experiments) and do not adequately reflect the generating causal function. Given an invariant causal structure, unstandardized regression coefficients are capable of detecting that invariance, whereas standardized regression coefficients are not. This is true even when the structural coefficients in the population are assumed to be standardized in form (see Kim & Ferree, 1981, for a more elaborate discussion of the issues involved). Thus, unstandardized coefficients are generally preferred to standardized ones and the DK transformation is limited, accordingly.

In sum, we recommend the centering approach advocated by Cronbach (1987), coupled with a formal evaluation (i.e., statistical test) of \( b_3 \) and an evaluation of \( b_1 \) at selected levels of \( X_2 \). The latter is analogous to simple main effects analysis in ANOVA frameworks and \( b_3 \) provides information corresponding to interaction comparison analysis. If standardized solutions are
sought, the DK transformation can be used in a similar fashion. However, this transformation is problematic where causal invariance is of concern.

*Reasons for Failures to Detect Interactions*

Additive transformations do not affect the overall test of statistical interaction, but they do affect the interpretability of regression coefficients. Similarly, contrary to Morris et al. (1986), multicollinearity probably is not the major culprit in psychologists’ failure to detect interaction effects in multiple regression analysis (barring computational errors). To the extent that multicollinearity is a problem, the transformations suggested in this article can mitigate its effects. However, we believe that such transformations will not be sufficient to adequately detect interaction effects. Rather, there are other reasons as to why interactions in multiple regression may be elusive. First, is the problem of measurement error. It is well known that unreliable measures can yield biased estimates of regression coefficients in multiple regression or structural equation analysis (e.g., Bohrnstedt & Carter, 1971). If measures are fallible, interaction effects may be difficult to detect (see Busemeyer & Jones, 1983, for elaboration). Second, is the issue of the functional form of the interaction. Traditional interaction analysis in multiple regression relies on a product term, $X_1 X_2$. This assesses interaction of a limited type, namely the case where the *linear* relationship between $Y$ and $X_1$ changes as a *linear* function of $X_2$. Thus, Equation 2 will be diagnostic of statistical interaction (as it is traditionally thought of by psychologists) only when the changes in the slope of $Y$ on $X_1$ are orderly and monotonic (linear) as one progresses across the levels of $X_2$. Given curvilinear relationships within levels of $X_2$, non-monotonic changes in slopes, or changes in slopes that are large when progressing between levels within one range of $X_2$,
but small when progressing between levels within a different range of $X_2$, the reliance on the traditional multiplicative term can, in many cases, be uninformative, if not misleading. Third, is the problem of levels of measurement. Although the evaluation of product terms is appropriate for interval level data, use of the approach on ordinal level data (as if the data had interval characteristics) may be problematic. Busemeyer and Jones (1983) present a convincing case for the biasing effects of departures from interval level data for the analysis of interaction effects. Finally, is the problem of statistical power. To the extent that analyses of interaction effects lack statistical power, the presence of the interaction is more likely to go undetected. Although it is difficult to specify the extent to which these four factors have contributed to a failure to observe expected interaction effects, they certainly represent issues that an investigator must consider when designing studies to explore interaction effects.

We conclude with the following recommendations:

1. To reduce potential problems with multicollinearity, the $X$ variables should be centered prior to the formation of product terms.

2. Given a statistically significant interaction effect that is bi-linear in form, perspectives on the nature of the interaction can be gained by application of the regression analogs of simple main effects analysis and interaction comparisons. These approaches explicitly recognize the conditional nature of slopes and standard errors. The centering transformation is readily amenable to execution of these analyses.

3. When concern is with causal analysis, unstandardized coefficients in interaction analysis are preferred to standardized coefficients because of their causal invariance. In this situation, the DK transformation is not recommended.

4. In the final analysis, the formulation and evaluation of interaction terms in multiple regression are best guided by a strong theory. The theory may suggest the optimal transformation for purposes of interpretation, the values of the moderator variable within which to pursue simple effects analysis, and/or the form of the interaction that is to be modeled. Wherever possible, theory should be brought to bear on statistical analyses of interactions.

5. Researchers should consider the potential effects of measurement error, levels of measurement, and statistical power when designing interaction studies using multiple regression.

References


Appendix

This appendix presents computational procedures for calculating \( \text{var}(b_1) \), \( \text{var}(b_2) \), \( \text{var}(b_3) \), \( \text{cov}(b_1,b_3) \), and \( \text{cov}(b_2,b_3) \) from standard computer output. Each regression coefficient in Equation 2, \( b_1 \), \( b_2 \), and \( b_3 \) will have an associated standard error. The \( \text{var}(b) \) for a given coefficient is the square of the standard error for that coefficient yielded by standard computer output. Let \( b_1' \) and \( b_2' \) represent the regression coefficients for \( X_1 \) and \( X_2 \) in Equation 1, respectively, and \( b_1 \), \( b_2 \), and \( b_3 \) represent the regression coefficients for \( X_1 \), \( X_2 \), and \( X_1X_2 \) in Equation 2, respectively. Let \( \text{MSE} \) represent the mean squared error from the overall analysis (usually reported in a summary table) for Equation 2 and \( \text{MSE}' \) represent the corresponding term for Equation 1. For analyzing the relationship between \( Y \) and \( X_1 \) at levels of \( X_2 \), the value of \( X_2 \) at which the minimal standard error occurs is

\[
\min X_2 = \frac{(b_1' - b_1)}{b_3}
\]

and

\[
\text{cov}(b_1,b_3) = \frac{(\text{MSE}/\text{MSE}') \text{var}(b_1') - \text{var}(b_1) - (\min X_2^2) \text{var}(b_3)}{2 \min X_2}
\]

The parallel formulas for analyzing the relationship between \( Y \) and \( X_2 \) at levels of \( X_1 \) are

\[
\min X_1 = \frac{(b_2' - b_2)}{b_3}
\]
and

\[
\text{cov}(b2, b3) = \frac{(\text{MSE'}/\text{MSE})\text{ var}(b2') - \text{var}(b2) - (\min X1^2)\text{ var}(b3)}{2 \min X1}
\]

All of the terms required to execute Equations 5 and 6 are thus defined. Some computer packages have options that permit the output of the above variances and covariances (e.g., the COVB option in SPSS-X REGRESSION).