JULIE R. IRWIN and GARY H. McCLELLAND*

Moderated multiple regression models allow the simple relationship between the dependent variable and an independent variable to depend on the level of another independent variable. The moderated relationship, often referred to as the interaction, is modeled by including a product term as an additional independent variable. Moderated relationships are central to marketing (e.g., Does the effect of promotion on sales depend on the market segment?). Multiple regression models not including a product term are widely used and well understood. The authors argue that researchers have derived from this simpler type of multiple regression several data analysis heuristics that, when inappropriately generalized to moderated multiple regression, can result in faulty interpretations of model coefficients and incorrect statistical analyses. Using theoretical arguments and constructed data sets, the authors describe these heuristics, discuss how they may easily be misapplied, and suggest some good practices for estimating, testing, and interpreting regression models that include moderated relationships.

Misleading Heuristics and Moderated Multiple Regression Models

A marketing researcher is interested in examining the effects of perceived quality and coupon availability (i.e., available versus not) on intent to purchase a product. The researcher expects both direct effects (intent will be greater for higher quality and when a coupon is present) and an interaction (the relationship between intent and quality will be stronger when a coupon is present). When analyzing the data, the researcher must decide what model to fit and how to interpret the estimated coefficients of that model. Suppose this marketing researcher decides to fit the following moderated regression model:[1]

(Q1) $\text{Intent} = a + b\text{Quality} + c\text{Coupon} + d(\text{Quality} \times \text{Coupon}).$

The following true/false quiz highlights important issues confronting the researcher who uses the moderated multiple regression model of Equation Q1:

*Julie R. Irwin is Assistant Professor of Marketing, University of Texas at Austin (e-mail: julie.irwin@bus.utexas.edu). Gary H. McClelland is Professor of Psychology, University of Colorado (e-mail: gary.mcclelland@colorado.edu). The authors thank Peter Fader, Paul Green, John Lynch, Carol Nickerson, Wesley Hutchinson, and the three anonymous *JMR* reviewers for their insightful and helpful comments. They also express their gratitude to Dilip Soman for generously discussing with them his data and analysis.

[1]All the models are assumed to have an error term. The error terms are omitted for notational simplicity, because our focus is the model and its parameters. Also, in several of our examples we use generic variables Y, X, and Z. The reader may note, however, that the quiz example maps directly onto these variables with Y = Intent, X = Quality, and Z = Coupon (a dichotomous variable indicating whether a coupon was available).

True/false:

1a. The relationship between Quality and Intent as assessed by coefficient b does not depend on whether the researcher uses dummy codes (0, 1 or 1, 0), contrast codes (–1, 1), or any other codes for Coupon. Similarly, the relationship between Coupon and Intent as assessed by coefficient c does not depend on how Quality is scaled.

1b. Changing the origin (i.e., subtracting or adding a constant) of either Quality or Coupon (or both) will change the test of the Quality × Coupon interaction as assessed by coefficient d.

2. A more accurate test of the interaction may result from testing the reduced model:

$$\text{(Q2)} \qquad \text{Intent} = a + d'(\text{Quality} \times \text{Coupon}),$$

because, for example, the coefficients b and c in Equation Q1 may not be statistically significant and/or both Quality and Coupon may be highly correlated with the product Quality × Coupon.

3. There would be no harm in converting the continuous independent variable Quality to a categorical variable (e.g., by splitting scores at the median into low and high quality groups) in order to use analysis of variance (ANOVA) and interpret the interaction using cell means.

Many researchers are surprised that, for reasons explained subsequently, the answer to all the quiz items is "false." A review of recent articles in marketing journals reveals that many authors believed that the answer to one or more of the quiz items was "true." This quiz presents some of the issues that are likely to arise in many marketing research projects in which moderator (interaction) relationships are expected. Moderated multiple regression augments the additive multiple regression model to allow the relationship between the dependent (criterion) variable and an independent (predictor) variable to depend on (i.e., be moderated by) the level of another independent variable. Polynomial regression models are a special case of moderated regression in which products (i.e., powers) of the same independent variable are included; such models allow the relationship between the independent and dependent variables to be moderated by the level of the same independent variable.

Multiple regression models not including a product term are widely used and well understood. Researchers can make understandable but serious mistakes when they apply heuristics they learned about from these simpler models to moderated multiple regression. This overgeneralization can happen for models estimated with ordinary least squares regression, analysis of covariance in which homogeneity of regression is tested, logistic regression, and structural equation modeling (LISREL, EQS, and their relatives). We use the term "moderated regression" for this entire class of models; our arguments hold equally for all of them. However, we explicitly do not consider multiplicative models in the context of information integration theory or axiomatic conjoint measurement, because those methods have their own procedures for resolving the issues we discuss here (see Lynch 1985).

We explain three of the heuristics that may be most easily misapplied and follow the explanations with short descriptions of good practices to use instead of these heuristics. Our goal is to present information that not only avoids incorrect interpretations but also leads to better understanding of the data. The proper use of moderated multiple regression is not a minor issue in marketing, but the issues we raise here have received little attention in the marketing literature. We were inspired to write this article when we noted that many of the heuristics appeared to underlie the testing and interpretation of moderated multiple regression models published recently in marketing and elsewhere. In nine recent issues of *Journal of Marketing Research* (Volumes 34 and 35, plus the first issue of Volume 36, the most recent issues when this article was written), of the 20 articles using moderated regression models, 18 (90%) applied one or more of the misleading heuristics described here. There were 15 (75%) examples of Heuristic 1, two (10%) examples of Heuristic 2, and 5 (25%) examples of Heuristic 3. The existence of even a single instance of these inappropriate heuristics is cause for concern about the methodological soundness of the conclusions reported for moderated regression models. We do not identify the authors of the articles, because our intent is not to single out those who have been deceived by these heuristics that on the surface appear quite reasonable.

Prior to the publication of this article, the canonical source for marketers performing moderated regression was by Sharma, Durand, and Gur-Arie (1981). This article presented a number of helpful regression tools, but we disagree with some of its conclusions. We respectfully suggest that our article serves, in part, as an updating and expansion of this earlier work.

We now turn to explanations of each heuristic and why it is inappropriate for moderated multiple regression. Mathematical details are contained in the Appendix.

HEURISTIC 1

In an additive regression model, changing the origin of one independent variable by adding or subtracting a constant has no effect on (1) the coefficients of the other independent variables, (2) the correlation or the covariance between the dependent variable and the rescaled independent variable, or (3) the correlation or the covariance between any two independent variables. For example, consider the additive regression model

$$\text{(1)} \qquad Y = a + bX + cZ,$$

in which X is a continuous independent variable (e.g., ratings of quality) and Z is a two-level categorical independent variable (such as whether a coupon was available). For this model, the heuristic is appropriate because mean centering (subtracting the mean of X from every value of X), standardizing,[2] or linearly rescaling X will not change the estimate of the partial regression slope c, its statistical test, or its interpretation. Similarly, changing the coding of Z to dummy codes (0, 1 or 1, 0), contrast codes (–1, 1 or 1, –1), or any other coding will not affect the estimate of the partial regression slope b, its statistical test, or its interpretation. Note that changing the scaling of X or the coding for Z will change the estimate of the intercept a, its test, and its interpretation, but the intercept often is not of substantive interest.

The heuristic about rescaling and recoding that is appropriate for additive multiple regression does not generalize to moderated regression models, which include interaction terms in the form of products of independent variables. As an example of where the heuristic fails, consider the addi-

[2]Standardized regression coefficients, or beta weights, are especially problematic in moderated models, because the z-score of a product does not equal the product of the z-scores (for an extended discussion of this issue, see Aiken and West 1991).

tion of the product $X \times Z$ to the model to yield this moderated regression model:

$$Y = a + bX + cZ + d(X \times Z). \tag{2}$$

We first consider the effect of rescaling the continuous independent variable X. We can factor Equation 2 in terms of a simple relationship between Z and Y; that is,

$$Y = (a + bX) + (c + dX)Z. \tag{3}$$

If Z is dummy-coded, then the slope $(c + dX)$ represents the difference between the two group means for Y at a particular level of X. The coefficient c depicts the difference between the two groups when and only when $X = 0$. In other words, c is the distance between the two regression lines at their intercepts. As X changes by one unit, the difference between groups will change by amount d. A common misinterpretation (perhaps as an overgeneralization from ANOVA) is to refer to c as the "main effect" of Z. The term "simple effect" is preferable, because this term refers to the simple relationship between the dependent variable and an independent variable at a particular level of the other independent variable(s).

Just as the intercept is the predicted value of the dependent variable when $X = 0$ in simple regression, the coefficient c is the predicted difference between groups at their intercepts (i.e., when $X = 0$). Thus, the location of the zero point matters, which violates the heuristic that changing the coding or scaling of one variable does not affect the estimation, testing, and interpretation of other slopes. This heuristic for additive multiple regression models does not apply to moderated regression models, which is why the answer to quiz item 1a is "false."

The intercept may have little or no meaning when $X = 0$ is outside the range of the independent variable; likewise, the coefficient c and its test may have little or no meaning when $X = 0$ is outside the range of the independent variable. If Quality in Equation Q1 were measured on a 1–7 scale, the Coupon coefficient would measure the effect of Coupon for a product with a nonexistent value for Quality.[3] It is easy to show that changing the origin of X will change the coefficient c but not the coefficient d. For example, suppose that X in Equation 2 is replaced in the model with $X' = X - \overline{X}$ so that the moderated regression model is

$$Y = a' + b'X' + c'Z + d'(X' \times Z), \tag{2'}$$

which factors to a model analogous to Equation 3,[4]

$$Y = (a' + b'X') + (c' + d'X')Z. \tag{3'}$$

Although Equation 2′ looks similar to Equation 2, their interpretations are different because the origin of X has changed. The coefficient c′ now estimates the difference between groups when $X' = 0$, or, equivalently, when X is at its average value. This is closer to what most researchers mean when they refer to the main effect. However, except for special situations (e.g., when the X distribution is exactly symmetric), it still will not be the same as the average difference between groups across all levels of X.

Many researchers indicate that a main effect was "changed" by the addition of the interaction, perhaps concluding that covariation with the product term reduced the power to detect the main effects. Sometimes the decision whether to include interaction terms is based on these "changes" in the pattern of effects. This is an unfortunate result of misunderstandings about how the addition of a moderator term influences the simple effects tests. An average or main effect does not change with the addition of a moderator term; the apparent change is the result of a shift from a main effect test to a test of a particular simple effect in the moderator model.

In Equation 3′, coefficient d′ describes how the difference between the groups changes for each one-unit increase in X. Given that the difference between the groups remains the same, no matter what the representation, it must be the case that $d = d'$. Hence, changing the origin of either variable does not affect the estimate of the interaction coefficient or its test, which explains why quiz item 1b is false.

An Illustrative Example

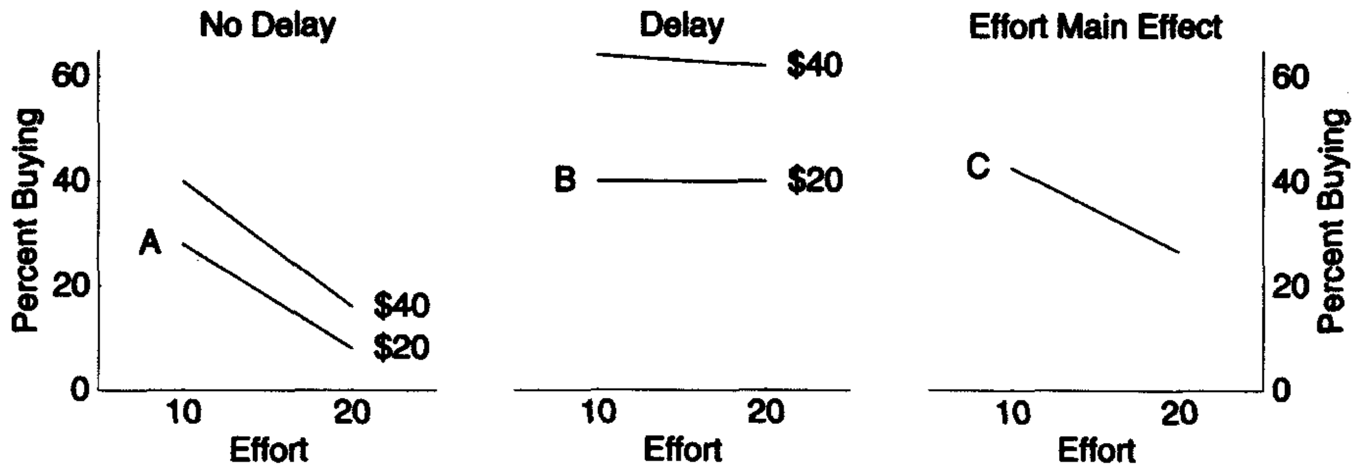
An example from a recent article published in *Journal of Marketing Research* illustrates the effects of coding changes on the statistical tests and interpretation of components of products in moderated multiple regression models. Figure 1, from Soman (1998), plots the percentages of consumers who indicate that they would purchase an item advertised with a rebate instead of a similar item without a rebate. The rebates varied in Delay (immediate versus delay of two or three weeks), Face Value ($20 versus $40 rebate for a $100 item), and Effort (needing to drive 10 versus 20 miles to a store to get the rebate). Soman (1998, p. 433) states, "The choice data were analyzed using a logistic regression model with … choosing the advertised brand as the dependent variable and dummies for delay, face value, and effort as the independent variables." However, Soman (1999) used the SAS procedure CATMOD, which by default uses contrast codes instead of dummy codes for the independent variables. Fortunately, all of Soman's (1998) interpretations are consistent with his actual analysis, but the results are inconsistent with a dummy code analysis. Considering the way different codings would have resulted in apparently different statistical results provides a concrete illustration that the invariance of tests in additive multiple regression when the other variables are recoded does not apply to moderated models.

Table 1 shows statistical tests for logistic regression models resulting from alternative codings of Soman's (1998) independent variables. The first column, labeled "Additive Model," presents results for an additive logistic regression model that does not include any product terms for the interactions; this column is invariant over the different codings. However, the subsequent three columns for moderated regression models including all the interactions do show differences due to coding. The first moderated model, labeled "Dummy Code 1," represents the dummy codes described but not used in the original article. The second moderated

[3]For a more extreme example, imagine a moderator model (such as Equation 2) in which Y is preference for apartments and independent variables X and Z are apartment size (in square feet) and number of windows, respectively. The X coefficient would then capture the effect of size for windowless houses, and the Z coefficient would capture the effect of windows for houses with zero square feet.

[4]Models with higher-order interactions do not factor in the easy way illustrated for interpreting models involving $X \times Z$; instead, to interpret slopes properly, partial derivatives of the moderated regression model must be taken (for details, see Aiken and West 1991; Judd and McClelland 1989).

Figure 1
PERCENTAGE OF CONSUMERS BUYING THE REBATED ITEM OVER THE REGULAR ITEM



Notes: Given as a function of Delay (the left-hand and center panels), Face Value (the two lines in the left-hand and center panels), and Effort (the end-points of each line). N = 25 for No Delay, N = 50 for Delay. Data are from Soman (1998).

Table 1
CHI-SQUARE VALUES FOR TERMS OF THE LOGISTIC REGRESSION MODEL USING ALTERNATIVE CODINGS

		Moderated Models		
	Additive Model	*Dummy Code 1*[a]	*Dummy Code 2*[b]	*Contrast Code*[c]
Delay (D)	21.88 (.0001)	1.03 (.31)	1.03 (.31)	21.07 (.0001)
Face Value (F)	11.45 (.0007)	.80 (.37)	5.65 (.02)	6.63 (.01)
Effort (E)	2.21 (.14)	3.02 (.08)	.00 (1.00)	5.21 (.022)
D × F	—	.36 (.55)	.36 (.54)	.20 (.66)
D × E	—	2.47 (.12)	2.47 (.12)	4.60 (.03)
F × E	—	.05 (.82)	.02 (.88)	.02 (.90)
D × F × E	—	.07 (.79)	.07 (.79)	.07 (.79)

[a]Immediate = 0, delay = 1; $20 face value = 0, $40 face value = 1; low effort = 0, high effort = 1.
[b]Immediate = 1, delay = 0; $20 face value = 0, $40 face value = 1; low effort = 0, high effort = 1.
[c]Immediate = –1, delay = 1; $20 face value = –1, $40 face value = 1; low effort = –1, high effort = 1.
Notes: Data are from Soman (1998). The values enclosed in parentheses are *p*-values.

model, labeled "Dummy Code 2," uses a dummy code for Delay that is reversed—the coding a researcher would have used if the variable had instead been called "Immediate," where 0 represents a delay and 1 represents an immediate rebate. The seemingly arbitrary choice of Delay versus Immediate coding affects the statistical tests of the other variables; for example, the chi-square value for Effort changes from 3.02 to .0. The last column, using contrast codes, depicts the results described by Soman (1998). The Delay variable is of particular interest in this experiment (because of the counterintuitive result that delaying the rebate increases purchase probability); a researcher using dummy codes who then mistakenly interprets the resulting chi-square test for Delay as a main effect would not have detected this interesting effect of Delay.

How can it be that different, but in all cases quite reasonable, coding schemes show such seemingly different results? The answer is that all three analyses describe the identical underlying model and are entirely consistent with one another; they simply highlight different aspects of the same model. Soman's regression equation is as follows:

$$(4)\ \text{Purchase} = a + b\text{Delay} + c\text{Face} + d\text{Effort} + e(\text{Delay} \times \text{Face}) + f(\text{Delay} \times \text{Effort}) + g(\text{Face} \times \text{Effort}) + h(\text{Delay} \times \text{Face} \times \text{Effort}),$$

which can be factored, as in Equation 2, to yield

$$(5)\quad \text{Purchase} = [a + b\text{Delay} + c\text{Face} + e(\text{Delay} \times \text{Face})] + [d + f\text{Delay} + g\text{Face} + h(\text{Delay} \times \text{Face})]\text{Effort}.$$

Equation 5 highlights tests of the effect of Effort on Purchase. As Equation 5 underscores, coefficients for components (in this case, d for Effort) in a moderated regression model reflect the simple relationship between that independent variable and the dependent variable when all other independent variables with which it is combined in interactions are equal to 0. That is, if both Delay and Face are 0, Equation 5 reduces to

$$(6)\qquad \text{Purchase} = a + d\text{Effort},$$

which represents the simple effect of Effort on Purchase when there is no delay (Delay = 0) and when face value equals \$20 (Face = 0). In the Dummy Code 1 analysis, the χ^2 value of 3.02 (p = .08) tests that simple effect. This Effort–Purchase simple relationship is represented by the line labeled "A" in the left-hand panel of Figure 1. The chi-square test of the coefficient d indicates that the percentages at either end of the line (28% and 8%) are not quite significantly different (p = .08).

In contrast, in the analysis for Dummy Code 2, the χ^2 value of 0 (p = 1.0) highlights the effect of Effort when there is a delay (Delay = 0 in this analysis using the reverse coding) and when the face value of the rebate is \$20 (Face = 0). This relationship is the line labeled "B" in the center panel of Figure 1; in this simple relationship, there is no effect of Effort. Coding Delay differently focuses the tests of Effort on different simple effects. Finally, using contrast codes tests the main effect of Effort averaged over the levels of both Delay and Face, because for both variables the mean contrast code is 0. For the main effect of Effort, the χ^2 value is 5.21 (p = .02) and is represented by the line labeled "C" in the right-hand panel of Figure 1. Thus, on average, Effort significantly decreased the percentage of respondents who would buy.

Note that the test of the highest-order interaction (Delay × Face × Effort, in this case) is invariant across alternative coding schemes. However, the tests of all the other components, including the lower-order interactions, are dependent on the coding scheme. All of these analyses are equally correct and legitimate. They simply highlight different simple effects in the same underlying model. Thus, researchers using any coding scheme would reach the identical conclusion about the highest-order interaction. The apparent differences in their conclusions about lower-order interactions and the individual components would be due to the different focus of each coding scheme. The lessons are (1) that heuristics from the additive model do not generalize to moderated models and (2) that knowing which coding schemes were used is therefore crucial to correct interpretation of the statistical analysis.

Good Practice 1a

Provide readers with detailed information about the coding of categorical independent variables and the origin or zero point of continuous independent variables and use this information to interpret coefficients and their tests. It is almost always a good idea to avoid using the term "main effect" in a regression model with an interactive term. Instead, coefficients of components of interactions represent the simple effects at particular levels of the other variables included in the interaction.

Good Practice 1b

Change the origin of each continuous independent variable and select the coding of categorical independent variables so as to focus the tests of the components of interactions on interesting theoretical or practical questions.

An Example of Judicious Coding

Good Practice 1b is illustrated in a recent *Journal of Marketing Research* article. LeClerc and Little (1997) investigate the effects of different kinds of advertising content on brand attitude for different levels of brand loyalty. One might have fit the following multiple regression model:

$$(7)\ \text{Brand Attitude} = a + b\text{Num} + c\text{Content} + d(\text{Num} \times \text{Content}),$$

where Num is the number of different brands the consumer had purchased in a given time period, and Content is a code indicating which of two advertising contents were used (an attractive picture versus brand information). With this coding, the parameter c for Content would have estimated the effect of the advertising content on brand attitude when Num = 0 (i.e., for people who have never bought any of the brands of this product). None of those people were included in the study, so c represents a prediction about cases that were of little interest. Instead, to focus the test on their main theoretical question, LeClerc and Little (1997) use Num′ = Num – 1 and estimate the model as follows:

$$(8)\qquad Y = a + b\text{Num}' + c\text{Content} + d(\text{Num}' \times \text{Content}).$$

In Equation 8, the coefficient c estimates the effect of advertising content when Num′ = 0. Requiring Num′ = 0 is the same as requiring Num = 1, which indicates the group of consumers who were highly brand loyal because they purchased only one of the possible brands at every opportunity. Thus, c assesses the effect of advertising content on brand attitude for the most loyal consumers. Carefully selecting the origin of the variable used in an analysis can change a coefficient (such as c in this example) from a meaningless estimate into a focused estimate of what is of greatest interest.

HEURISTIC 2

Variables can be added or deleted to the regression model depending on their contribution to the overall fit and/or whether the researcher has specific hypotheses about the variables.

Sometimes researchers choose to include independent variables on the basis of their contribution to model fit. This criterion underlies the use of stepwise regression and other formal and informal variable selection methods. Many data analysts question the advisability of these methods even for additive multiple regression (e.g., Henderson and Velleman 1981). However, stepwise regression and the more general practice of deciding whether to include or exclude variables in a final model on the basis of fit are sufficiently common to warrant careful examination of whether this heuristic applies to moderated multiple regression. Even when researchers are not explicitly using stepwise regression, they may decide to eliminate variables from a model because of lack of significance or lack of interest in those variables. For example, a researcher may be primarily interested in an interactive effect and not care much about the average effect.

Applying Heuristic 2 to models involving product variables, researchers sometimes compare (explicitly or implicitly) the following two models:

(9) Product Model: $Y = a' + d'(X \times Z)$

and

(10) Moderated Model: $Y = a + bX + cZ + d(X \times Z)$.

The comparison of these models is a generalization of the common practice of comparing the contribution of an independent variable to model fit by testing the model with and without the variable. In this case, the comparison is between the moderated model with the simple effects and the product model without the simple effects. Not infrequently, the product term is significant in the product model (Equation 9), but not in the moderated model (Equation 10). And especially when the components X and Z themselves are not significant in the moderated model (or are not of interest), researchers opt for the product model. Some important articles on moderated multiple regression (e.g., Sharma, Durand, and Gur-Arie 1981) have recommended models that did not include all component terms.

This strategy and the heuristic that underlies it are inappropriate for moderated multiple regression, because the correlation between Y and $X \times Z$ is affected by changes in the origin of either X or Z. To illustrate this fact as well as the next heuristic, we use the hypothetical data presented in Table 2. The columns represent independent variables X (continuous) and Z (dichotomous) and dependent variable Y. The correlation between Y and $X \times Z$ is .96; however, if mean centering is used instead so that $X' = X - 4$, then the correlation between Y and $X' \times Z$ drops to .54. Figure 2 depicts the correlation between Y and the product of the two independent variables across a range of changes of origin for X (i.e., across values of $X - k$, where k is a constant). The correlation is zero when $k = 16/3$ (i.e., when $X' = X - 16/3$) and ranges between −.81 and +.97. Decisions about the product model would depend solely on which origin the researcher used; this is clearly unacceptable. Therefore, quiz item 2 is false. In contrast, note that the partial correlation between Y and $X' \times Z$, controlling for X' and Z, remains constant at .38 across changes of origin for X. The partial correlation between Y and the product $X \times Z$, controlling for X and Z, is the relevant statistic for evaluating the product component. To ensure that the components are partialed from the product, both components X and Z must be included in the model whenever the product $X \times Z$ is included.

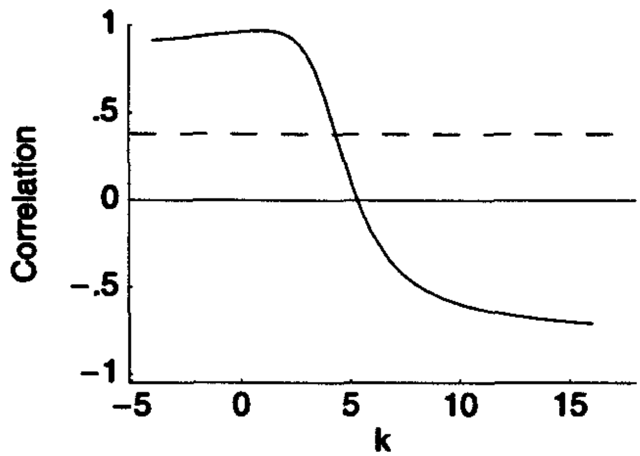
Table 2

HYPOTHETICAL DATA FOR ILLUSTRATING MODERATED MULTIPLE REGRESSION ANALYSIS

Observation	*Y*	*X*	*Z*
1	5	2	0
2	3	3	0
3	4	4	0
4	3	5	0
5	5	6	0
6	7	2	1
7	7	3	1
8	10	4	1
9	11	5	1
10	15	6	1

Figure 2

CORRELATION (SOLID LINE) BETWEEN $(X - k) \times Z$ AND Y AND PARTIAL CORRELATION (DASHED LINE) BETWEEN $(X - k) \times Z$ AND Y CONTROLLING FOR $(X - k)$ AND Z



Some researchers omit a component independent variable from the moderated regression model after noting that the statistical significance of the product term is increased when the component variable is not in the model. If a component of the product is eliminated, the test of the product is confounded with the main effect of that component independent variable on the dependent variable. It should not be surprising that a regression term confounding two effects would have a greater likelihood of being statistically significant, especially if the two confounded effects were in the same direction. However, the confounding makes the test of the product term uninterpretable. It is only when all the components of a product term are included in the model that the product term represents the interaction (Cohen 1978; Cronbach 1987).

Good Practice 2

Whenever a product term is included in the model, be sure to include all of its components. For stepwise regression, order of entry must be controlled so that the product never enters before any of its components, and if the product is in the model, none of the components should be deleted. Even if components are not significant, they must remain in the moderated regression model to enable the proper partialling of the product.

HEURISTIC 3

To aid in interpretation and avoid scaling problems when interactions are tested, split continuous variables into groups so that ANOVA can be used.

Median splits, common in marketing, are motivated to facilitate both interpretation of interactions and their statistical tests. Personality and segmentation variables, in particular, often are split to form groups so that ANOVA can be used. There are intriguing measurement and theoretical issues underlying the choice to split variables into groups, but here we focus only on the statistical implications of median splits in moderated models.

The explication of an interaction seems more difficult with continuous than with discrete independent variables: Although the regression coefficient of the interaction term may be significant, what does that coefficient mean? Until recently, few textbooks included explanations about how to test and interpret interactions involving continuous variables in moderated regression models (for exceptions, see Aiken and West 1991; Jaccard, Turrisi, and Wan 1990; Judd and McClelland 1989). Splitting the continuous independent variables into two or more groups may be appealing to those trained in an ANOVA framework because the interaction appears easier to explicate. Unfortunately, median splits are likely to change the statistical significance of the interaction.

Again, consider the illustrative hypothetical data of Table 2. Table 3 presents three moderated regression models. The first is for the untransformed variables of Table 2; the t-statistic for the interaction is 3.87 ($p = .01$). The next two columns show analyses for two alternative median splits for X. In "Median Split High," values of X exactly equal to the median are placed in the high group, whereas in "Median Split Low," they are placed in the low group. In both analyses, the low group is coded as 0 and the high group as 1. Using either median split, the t-statistic for the interaction drops to 2.24 ($p = .07$) and is no longer statistically significant at conventional levels. In other words, dichotomizing decreased the statistical power for detecting the interaction. To have produced the same t-statistic as the original analysis, the median split analyses would require approximately 21 observations, slightly more than twice the original number of observations. Many articles (e.g., Aiken and West 1991; Cohen 1983) demonstrate that for bivariate normal distributions and uncorrelated independent variables, median splits can severely reduce statistical power. Dichotomizing a variable with a median split can be equivalent to discarding approximately half of the data (McClelland 1997). Given that power for detecting interactions with continuous variables is usually low (McClelland and Judd 1993; Stone-Romero, Alliger, and Aguinis 1994), adopting an analysis strategy that reduces power still further by decreasing the effective sample size by half makes no sense.

Surprisingly, median splits can induce spurious interactions when the variables are correlated (Maxwell and Delaney 1993). Splitting data into groups can confound curvilinear trends in the interaction components with interactions between them, resulting in significant interactions when there actually is no interaction present. Thus, quiz item 3 is false; median splits are not a good practice because they alter significance patterns and can increase the possibility of both Type I and Type II errors.

Table 3

REGRESSION COEFFICIENTS AFTER TWO DIFFERENT MEDIAN SPLITS OF VARIABLE X

	$X' = X$ $Z' = Z$	X' = *Median Split High*[a] $Z' = Z$	X' = *Median Split Low*[b] $Z' = Z$
Intercept	4	4	4
	(2.58)	(3.27)	(4)
X′	0	0	0
	(0)	(0)	(0)
Z′	–2	3	4
	(–.91)	(1.73)	(2.83)
X′Z′	2	5	5
	(3.87)	(2.24)	(2.24)

[a]Observations below, above, and equal to the median are coded 0, 1, and 1, respectively.

[b]Observations below, above, and equal to the median are coded 0, 1, and 0, respectively.

Notes: The analyses are for the hypothetical data in Table 2. Z is coded {0, 1} as shown in Table 2. The values enclosed in parentheses are t-statistics.

Good Practice 3

Do not reduce independent variables to a smaller number of categories as is done in median splits.

A Note on the Interpretation of Interactions

As mentioned previously, some researchers are motivated to split continuous independent variables into groups because they believe doing so will simplify the explication of interactive effects. An interaction can be understood as changes in the effect of one independent variable across levels of another. To explicate the interaction, the researcher needs to know how the simple effects change across levels of the other (moderator) variable. If one variable is discrete, as in Soman's (1998) data, then this explication is fairly straightforward: Examine the slopes of the continuous independent variable across the different levels of the discrete independent variable. We examined the variable Effort in this way previously in this article, using the information provided in Figure 1 and Table 1.

If both independent variables are continuous, the researcher can examine the differential effects of one independent variable across values of the other (Aiken and West 1991; Jaccard, Turrisi, and Wan 1990). This practice is different from collapsing the data into groups; the data remain in their original form, but a "spotlight" is focused on the model from different angles. Whether one or both independent variables are continuous, we recommend probing the interaction using the principles underlying the explication of Heuristic 1. The researcher simply changes the level of the moderator variable to understand the interactive effects on the independent variable of interest. Which variable is designated the moderator is arbitrary.

Suppose a researcher would like to probe the data from Table 2. Assume Y is intent to purchase, X is a quality rating, and Z is the presence or absence of a coupon. If effects of X on Y are of primary interest, then substituting different values of Z in the moderated multiple regression equation produces separate equations for each level of Z. For example, the moderated regression model using the dummy coding as in the original data of Table 2 is

(11) $Y = 4 + 0X - 2Z + 2(X \times Z)$.

Substituting 0 for Z (corresponding to no coupon) yields

(12) $Y = 4 + 0X$.

And substituting 1 for Z (corresponding to having a coupon) yields

(13) $Y = 2 + 2X$.

The two separate equations are plotted in the left-hand side of Figure 3. The results of the interaction are now easy to describe. On the one hand, there is no relationship (i.e., the slope is 0) between Quality (X) and Intent (Y) when there is no coupon ($Z = 0$). On the other hand, there is a positive relationship between X and Y when there is a coupon ($Z = 1$). Note that the interaction has already been shown to be significant; the spotlight method merely reveals the nature of the interaction.

Altering the coding for Z tests the two simple effects of X on Y, represented by the slope for X in Equations 12 and 13. The original coding (0 = no coupon and 1 = coupon) tests whether the slope of 0 in Equation 12 is significantly different from zero (obviously not, in this case). An alternative coding for Z for which $Z = 0$ represents having the coupon (such as the reverse dummy codes 1 = no coupon and 0 = coupon) would test whether the slope of 2 for X in Equation 13 is significantly different from zero.

The same explication strategy is effective for specifying the relationship between Coupon (Z) and Intent (Y) at different levels of the continuous variable Quality (X). Select several meaningful values for X, such as the mean Quality of ($X = 4$), a low level of Quality (say, two units below the mean, or $X = 2$), and a high level of Quality (say, two units above the mean, or $X = 6$). Substituting these values into Equation 11 yields[5]

(14) $Y = 4 + 2Z$ when $X = 2$ (low Quality),

$Y = 4 + 6Z$ when $X = 4$ (average Quality), and

$Y = 4 + 10Z$ when $X = 6$ (high Quality).

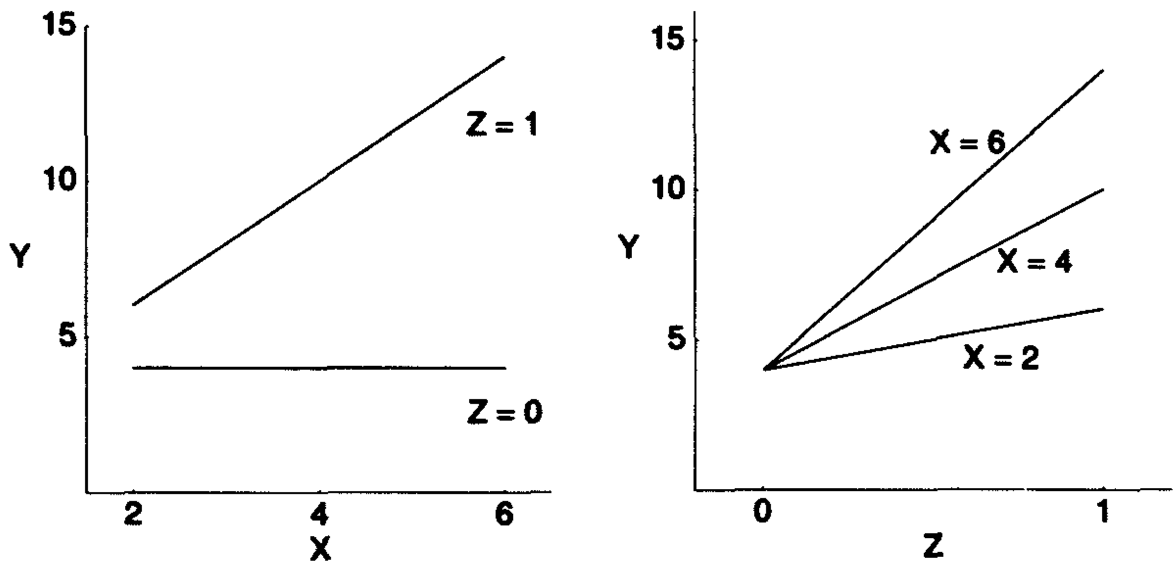
These three regression models are plotted in the right-hand side of Figure 3. Remember that the slope for the dummy-coded variable Z represents the mean difference in Intent for having the coupon versus not having it. Therefore, Equation 14 shows that the coupon increases intent to purchase by 2 when quality is low, by 6 when quality is average, and by 10 when quality is high. The effect of the coupon increases as quality increases. This method of interpreting interactions is similar to that described by Aiken and West (1991). To test whether the three differences or simple effects are significantly different from zero, simply test the coefficient for Z in three new moderated multiple regression models in which X is replaced, respectively, by $X - 2$, $X - 4$, and $X - 6$.

GENERAL CONCLUSION

Most marketing research projects, whether academic or industrial, will investigate questions involving moderated multiple regression models. Marketers expect that attributes may be weighted differently by different segments, that the influence of advertisements may depend on degree of

[5]Any moderated multiple regression model (Equation 11 or models resulting from using any linear recoding of the variables) for the data of Table 2 makes the same predictions. The fundamental underlying model is always identical. Care should be taken that the values selected for substitution are consistent with the variable scaling.

Figure 3
EXPLICATION OF THE INTERACTION OF TABLE 2



Notes: Left-hand panel depicts simple relationships between Quality (X) and Intent (Y) with and without Coupon (Z); right-hand panel depicts simple relationships between Coupon (Z) and Intent (Y) at different levels of Quality (X).

involvement, that promotions may be more effective the more price sensitive a shopper is, and so on. In approaching these questions of independent variable X moderating the relationship between independent variable Z and dependent variable Y, it is natural to apply some of the many heuristics, rules, and words of wisdom regarding additive multiple regression and ANOVA. However, these heuristics often do not generalize to moderated multiple regression. As we have shown, including a product or power term in a model changes the analysis and its interpretation in fundamental ways; incomplete understanding of these changes undoubtedly has led to some confusion in the literature. Following the good practices recommended here should facilitate the testing, interpretation, and explication of moderated multiple regression models.

APPENDIX

It is well known that in additive regression models, changing the origin of one independent variable by adding or subtracting a constant has no effect on (1) the coefficients of the other independent variables, (2) the correlation or the covariance between the dependent variable and the rescaled independent variable, or (3) the correlation or the covariance between any two independent variables. For example, in the additive regression model

$$Y = a + bX + cZ, \tag{A1}$$

changing the origin of Z has no effect on the coefficient b of X, the correlation or the covariance between Y and X or between Y and Z, or the correlation or the covariance between X and Z. None of these three invariances holds for the moderated regression model

$$Y = a + bX + cZ + d(X \times Z). \tag{A2}$$

If there is a nonzero interaction (i.e., $d \neq 0$ in Equation A2), then changing the origin of Z has the following effects:

1. There is a change in the coefficient b of X, and there exists a change of origin for Z such that b equals zero.
2. There is a change in the covariance between Y and XZ, and there exists a change of origin for Z such that this covariance equals zero.
3. There is a change in the covariance between X and XZ, and there exists a change of origin for Z such that this covariance equals zero.

Effect of Rescaling Z on the Coefficient for X

Suppose that a constant k is subtracted from Z so that

$$Z' = Z - k, \tag{A3}$$

which is equivalent to

$$Z = Z' + k. \tag{A4}$$

Substituting this value for Z into the moderated multiple regression Equation A2 yields

$$Y = a + bX + c(Z' + k) + dX(Z' + k). \tag{A5}$$

Expanding and grouping terms gives the equivalent model:

$$Y = (a + ck) + (b + dk)X + cZ' + dXZ'. \tag{A6}$$

In the original moderated regression model in Equation A2, the intercept equals a. After rescaling Z, the intercept equals (a + ck). Rescaling Z in the additive regression model in Equation A1 also results in this change in the intercept. The coefficient of X has also changed; it equals b in Equation A2 and (b + dk) in Equation A6. However, the coefficients of Z′ and XZ′ in Equation A6 equal the coefficients of Z and XZ, respectively, in Equation A2. That is, the coefficients of Z (c) and XZ (d) do not change even though Z has been rescaled. Thus, changing the origin of one component of the product term in a moderated regression model changes the coefficient of the other component of the product term and the intercept; it does not change the coefficient of the rescaled term or the coefficient of the product term.

If there is a nonzero interaction (i.e., $d \neq 0$ in Equation A2), then the change of origin $k = -b/d$ yields the model

$$Y = (a - bc/d) + cZ' + dXZ', \tag{A7}$$

in which the coefficient of X equals zero.

Effect of Rescaling Z on the Covariance Between Y and XZ

Bohrnstedt and Goldberger (1969) prove that the covariance of two products WY and XZ equals

$$\begin{aligned} CV[WY,XZ] = {} & -CV[W,Y]\,CV[X,Z] \\ & + CV[Y,Z]\,EV[W]\,EV[X] \\ & + CV[W,Z]\,EV[X]\,EV[Y] \\ & + CV[Y,X]\,EV[W]\,EV[Z] \\ & + CV[W,X]\,EV[Y]\,EV[Z] \\ & + EV[Z]\,EV[dv[W]\,dv[X]\,dv[Y]] \\ & + EV[Y]\,EV[dv[W]\,dv[X]\,dv[Y]] \\ & + EV[X]\,EV[dv[W]\,dv[Y]\,dv[Z]] \\ & + EV[W]\,EV[dv[X]\,dv[Y]\,dv[Z]] \\ & + EV[dv[W]\,dv[X]\,dv[Y]\,dv[Z]], \end{aligned} \tag{A8}$$

where

CV = the covariance between two variables,
EV = the expected value (mean) of a variable, and
dv = the deviation of a variable from its mean.

The covariance between Y and XZ can be found by substituting the constant 1 for W in Equation A8. That is,

$$\begin{aligned} CV[Y, XZ] = {} & CV[Y,Z]\,EV[X] + CV[Y,X]\,EV[Z] \\ & + EV[dv[X]\,dv[Y]\,dv[Z]]. \end{aligned} \tag{A9}$$

Unlike the covariance between Y and X or between Y and Z in the additive regression model, the covariance between Y and XZ (CV[Y,XZ]) in the moderated regression model depends on the expected values or means of X and Z. Changing the origin of X or Z or both thus necessarily changes the covariance between Y and XZ. In particular, if CV[Y,X] does not equal zero, then changing the origin of Z so that

$$CV[Z] = \frac{CV[Y,Z]\,EV[X] + EV[dv[X]\,dv[Y]\,dv[Z]]}{CV[Y,X]} \tag{A10}$$

guarantees that the covariance between Y and XZ equals zero. (If CV[Y,X] equals zero but CV[Y,Z] does not equal zero, then a similar change of origin for X ensures that

CV[Y,XZ] equals zero. In the unlikely event that both CV[Y,Z] and CV[X,Z] equal zero, then CV[Y,XZ] is invariant under changes of origin and is equivalent to EV[dv[X] dv[Y] dv[Z]]). Hence, most tests of the product model

$$\text{(A11)} \qquad Y = a + dXZ$$

depend on the arbitrary choice of origins for X and Z. As Holbrook (1977), Cohen (1978), Cronbach (1987), and others have noted, this arbitrariness is eliminated if the components of a product term are included in the regression model regardless of whether their coefficients are significantly different from zero. Then, the coefficient d for the product term is independent of changes of origin for either X or Z or both.

Effect of Rescaling Z on the Covariance Between X and XZ

Similarly, the covariance between X and XZ can be found by substituting X for W and the constant 1 for Y in Equation A8:

$$\text{(A12)} \quad CV[X,XZ] = CV[X,Z]\,EV[X] + EV[dv[X]^2\,dv[Z]] + EV[Z]\,V[X],$$

where V[X] = the variance of X.

Unlike the covariance between X and Z in the additive regression model, the covariance between X and XZ (CV[X,XZ]) in the moderated regression model depends on the expected values or means of X and Z. Changing the origin of X or Z or both thus necessarily changes the covariance between X and XZ. In particular, changing the origin of Z so that

$$\text{(A13)} \quad EV[Z] = -\frac{CV[X,Z]\,EV[X] + EV[dv[X]^2\,dv[Z]]}{V[X]}$$

guarantees that the covariance between X and XZ equals zero. Therefore, any collinearity between X and XZ can always be eliminated by some change in the origin of Z. If both X and Z are symmetrically distributed, then the last expected value in the numerator of Equation A13 equals zero. Then, mean centering (i.e., subtracting the value of the mean of a variable from each value of the variable) eliminates the term CV[X,Z] EV[X] in Equations A12 and A13. Thus, mean centering both X and Z eliminates the covariance between either component of the product term and the product term itself.

If X and Z are jointly skewed, then mean centering ensures that

$$\text{(A14)} \qquad CV[X,XZ] = EV[dv[X]^2\,dv[Z]].$$

Hence, after mean centering, the covariance between X and XZ is very small unless X and Z are extremely jointly skewed.

REFERENCES

Aiken, Leona S. and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage Publications.

Bohrnstedt, G.W. and A.S. Goldberger (1969), "On the Exact Covariance of Products of Random Variables," *Journal of the American Statistical Association*, 64 (328), 325–28.

Cohen, Jacob (1978), "Partialed Products *Are* Interactions; Partialed Powers *Are* Curve Components," *Psychological Bulletin*, 85 (4), 858–66.

——— (1983), "The Cost of Dichotomization," *Applied Psychological Measurement*, 7 (3), 249–53.

Cronbach, Lee J. (1987), "Statistical Tests for Moderator Variables: Flaws in Analysis Recently Proposed," *Psychological Bulletin*, 102 (3), 414–17.

Henderson, H. and Paul F. Velleman (1981), "Building Multiple Regression Models Interactively," *Biometrics*, 37 (2), 391–411.

Holbrook, Morris (1977), "Comparing Multiattribute Models by Optimal Scaling," *Journal of Consumer Research*, 4 (3), 165–71.

Jaccard, James, R. Turrisi, and C.K. Wan (1990), *Interaction Effects in Multiple Regression*. Newbury Park, CA: Sage Publications.

Judd, Charles M. and Gary H. McClelland (1989), *Data Analysis: A Model Comparison Approach*. San Diego: Harcourt Brace Jovanovich.

Leclerc, France and John D.C. Little (1997), "Can Advertising Copy Make FSI Coupons More Effective?" *Journal of Marketing Research*, 34 (4), 473–84.

Lynch, John G., Jr. (1985), "Uniqueness Issues in the Decompositional Modeling of Multiattribute Overall Evaluations: An Information Integration Perspective," *Journal of Marketing Research*, 22 (1), 1–19.

Maxwell, Scott E. and Harold D. Delaney (1993), "Bivariate Median Splits and Spurious Statistical Significance," *Psychological Bulletin*, 113 (1), 181–90.

McClelland, Gary H. (1997), "Optimal Design in Psychological Research," *Psychological Methods*, 2 (1), 3–19.

——— and Charles M. Judd (1993), "Statistical Difficulties of Detecting Interactions and Moderator Effects," *Psychological Bulletin*, 114 (2), 376–90.

Sharma, Subhash, Richard M. Durand, and Oded Gur-Arie (1981), "Identification and Analysis of Moderator Variables," *Journal of Marketing Research*, 18 (3), 291–300.

Soman, Dilip (1998), "The Illusion of Delayed Incentives: Evaluating Future Effort–Money Transactions," *Journal of Marketing Research*, 35 (4), 427–37.

——— (1999), personal e-mail communication, (March 16).

Stone-Romero, Eugene F., George M. Alliger, and Herman Aguinis (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 (1), 167–78.